R E P O R T R E S U M E S

ED 017 909 AL 001 091

THE PHONOLOGICAL SYSTEM OF CANTONESE.

BY- CHENG, TERESA
CALIFORNIA UNIV., BERKELEY

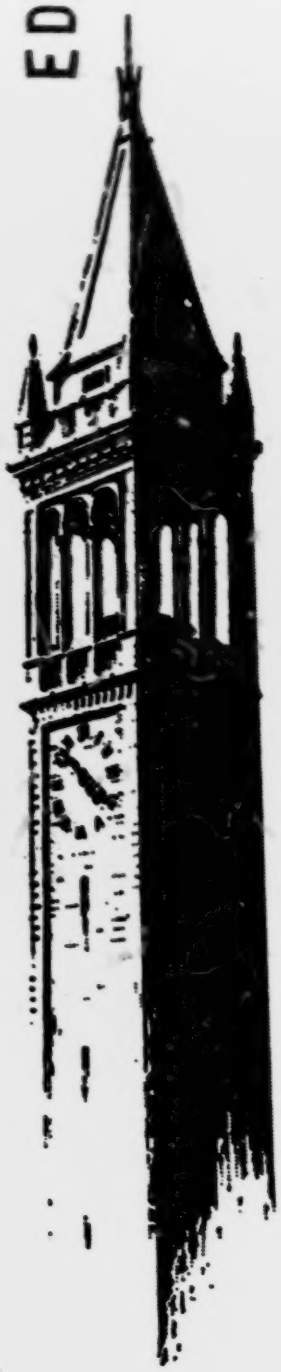
REPORT NUMBER POLA-2-5-1 PUB DATE MAR 68
NATIONAL SCIENCE FOUNDATION, WASHINGTON, D.C.
REPORT NUMBER GS-1430
EDRS PRICE MF-$0.50 HC-$3.84 94P.

DESCRIPTORS- *PHONOLOGY, *CANTONESE, TONE LANGUAGES, DIALECTS, LANGUAGE RESEARCH, SYLLABLES, DISTINCTIVE FEATURES, LINGUISTIC THEORY, PHONEMICS, GRAMMAR,

THIS STUDY OF THE CANTONESE PHONOLOGICAL SYSTEM IS A CASE STUDY OF THE PHONOLOGICAL COMPONENT OF A GENERATIVE GRAMMAR. ATTEMPTS ARE MADE IN CHAPTER III TO SOLVE SOME OLD PROBLEMS CONNECTED WITH THE ANALYSIS OF CANTONESE WITHIN THIS NEW THEORETICAL FRAMEWORK. MOST OF THE PROBLEMS ARE PSEUDO-PROBLEMS, IN THE SENSE THAT THEY POSE DIFFICULTIES ONLY FOR "DISTINCTIVE SEGMENT ANALYSIS" (THE AUTHOR'S TERM FOR PHONEMIC ANALYSIS), BUT NOT FOR THIS DISTINCTIVE-FEATURE ANALYSIS. THE PHONOLOGICAL SYSTEM OF CANTONESE IS PRESENTED IN CHAPTER II, IN WHICH THE SOUNDS ARE GIVEN DISTINCTIVE FEATURE REPRESENTATIONS, AND IN CHAPTER IV, IN WHICH SYLLABLE STRUCTURE IS STUDIED. SEQUENTIAL CONSTRAINTS AMONG SOUNDS ARE USED AS A CUE TO SIMPLIFY LEXICAL OR SYSTEMATIC PHONEMIC REPRESENTATIONS OF CANTONESE WORDS. ADDITIONAL INFORMATION CONCERNING THIS DOCUMENT MAY BE OBTAINED BY WRITING TO WILLIAM S-Y. WANG, DEPARTMENT OF LINGUISTICS, UNIVERSITY OF CALIFORNIA, BERKELEY, CALIFORNIA 94720. (AUTHOR/AMM)

PHONOLOGY LABORATORY - DEPARTMENT OF LINGUISTICS
UNIVERSITY OF CALIFORNIA, BERKELEY

ED017909



Project On Linguistic Analysis

Reports. Second Series, No. 5: March 1968

Teresa Cheng. The Phonological System of Cantonese C1-C77
Chao-hui Tung. The Phonological System of Gaoxiong, a Min Dialect of Chinese T1-T85

AL 001 091

Reproduction in whole or in part is permitted for any purpose of the United States Government.

The Project on Linguistic Analysis is supported in part by the National Science Foundation (Grant GS1430), the Office of Naval Research (Contract N00014-67-A-0114-0005), and the Air Force (Contract F30602-67-C-0347). It is administered through the Phonology Laboratory of the University of California at Berkeley, which has its office in 51 Dwinelle Hall (Telephone: 845-6000, Extension 1507).

Correspondence concerning POLA matters should be addressed to William S-Y. Wang, Department of Linguistics, University of California, Berkeley, California 94720.

Papers which have appeared in POLA, Second Series (as these are published, bibliographical references will be added):

No. 1: September, 1967

Introduction
Report of the Phonology Laboratory: May, 1967
Report of the Phonology Laboratory: August, 1967
Chin-chuan Cheng. English stresses and Chinese tones in Chinese sentences (Phonetica).
Burckhard Mohr and William S-Y. Wang. Perceptual distance and the specification of phonological features (Phonetica).
Marilyn May Vihman. Palatalization in Russian and Estonian.

No. 2: October, 1967

Wallace L. Chafe. The ordering of phonological rules (IJAL).
William S-Y. Wang. Competing changes as a cause of residue (Language).
Haruo Aoki. Toward a typology of vowel harmony (IJAL).
C. Douglas Johnson. A note on precedence relations in phonology.

No. 3: November, 1967

William S-Y. Wang. Conjoining and deletion in Mandarin syntax (Monumenta Serica).
Robert L. Cheng. Universe-scope relations and Mandarin noun phrases.

No. 4: January, 1968

William S-Y. Wang. The basis of speech.

No. 5: March, 1968

Teresa Cheng. The phonological system of Cantonese.
Chao-hui Tung. The phonological system of Gaoxiong, a Min dialect of Chinese.

The Phonological System of Cantonese*

Teresa Cheng

University of Hawaii

U.S. DEPARTMENT OF HEALTH, EDUCATION & WELFARE
OFFICE OF EDUCATION

THIS DOCUMENT HAS BEEN REPRODUCED EXACTLY AS RECEIVED FROM THE PERSON OR ORGANIZATION ORIGINATING IT. POINTS OF VIEW OR OPINIONS STATED DO NOT NECESSARILY REPRESENT OFFICIAL OFFICE OF EDUCATION POSITION OR POLICY.

"PERMISSION TO REPRODUCE THIS [redacted] MATERIAL HAS BEEN GRANTED BY Public domain

TO ERIC AND ORGANIZATIONS OPERATING UNDER AGREEMENTS WITH THE U.S. OFFICE OF EDUCATION. FURTHER REPRODUCTION OUTSIDE THE ERIC SYSTEM REQUIRES PERMISSION OF THE [redacted] OWNER."

*This study was supported by National Science Foundation Grant GS1430.

AL 001 091

Table of Contents

List of Figures 1

Abstract 2

Dialect maps of China 3

Chapter I. Introduction 6

1.1 The Cantonese dialect 6

1.2 Theoretical background of analysis 6

Chapter II. The sounds of standard Cantonese 10

2.1 Major class features of Cantonese segmental sounds 10

2.2 The segmental system of Cantonese 11

2.3 Redundancy rules 13

2.4 The tonal system of Cantonese 15

2.5 Redundancy rules 16

Chapter III. Problems in the treatment of Cantonese phonology 18

3.1 Some deviations from standard Cantonese 18

3.2 Initials 19

3.3 Medials 24

3.4 Finals 26

3.5 Reading (wén) vs. colloquial (bái) pronunciation 29

3.6 The alternation between the yin ping tones ㄱ (55) and ㄴ (53) 33

Chapter IV. Syllable structure 41

4.1 The canonical form of Cantonese syllables 41

4.2 Some general facts about Cantonese syllable structure 42

4.3 The distributional pattern of Cantonese sounds 43

4.4 Sequential constraints 46

References 80

Footnotes 81

Numerical 81

Alphabetical 82

List of Figures

Figure Ia.* Distribution of the major dialects of China
(A) Mandarin; (B) Wú; (C) Mǐn; (C') Southern Mǐn;
(D) Hakka; (E) Yùe 3

Figure Ib. The five subgroups of Yùe dialects
I. Yùe hǎi; II. Sì yì; III. Gao-Léi; IV. Qin-Lián;
V. Gùi nán 4

Figure Ic. Distribution of the Yùe dialects 5

Figure II. A model of sound-meaning correlations 7

Figure III. Natural classification of the nasal consonants 10

Figure IV. Major classes of Cantonese sounds 11

Figure Va. The segmental system of Cantonese 14

Figure Vb. The segmental system of Cantonese: features generated
by redundancy rules 14a

Figure VIa. The tonal system of Cantonese 17

Figure VIb. The tonal system of Cantonese: features generated by
redundancy rules 17a

Figure VII. Long rù finals 27

Figure VIII. Short rù finals 28

Figure IX. Cantonese finals 44

Figure X. Medial-final combinations on the systematic phonemic
level 45

*Figure Ia is a modification of Map 18 (Major languages) in Chi-yun Cheng's National atlas of China: dialect boundaries not relevant to the present study have been eliminated. Figures Ib and Ic are my enlargements of the shaded area in Figure Ia.

List of Figures (Cont.)

Figure XI. Medial-final combinations on the systematic phonetic level 46

Figure XII. The total distributional pattern 47

Abstract

This study of the Cantonese phonological system is, in fact, a case study of the phonological component of a generative grammar. Attempts are made in Chapter III to solve some old problems connected with the analysis of Cantonese within this new theoretical framework. It turns out that most problems are pseudo- problems, in the sense that they pose difficulties only for 'distinctive segment analysis' (my term for phonemic analysis), but not for our distinctive-feature analysis. The phonological system of Cantonese is presented in Chapters II and IV. In Chapter II, Cantonese sounds are given distinctive-feature representations. In Chapter IV, syllable structure is studied. Sequential constraints among sounds are used as a cue to simplify lexical or systematic phonemic representations of Cantonese words.

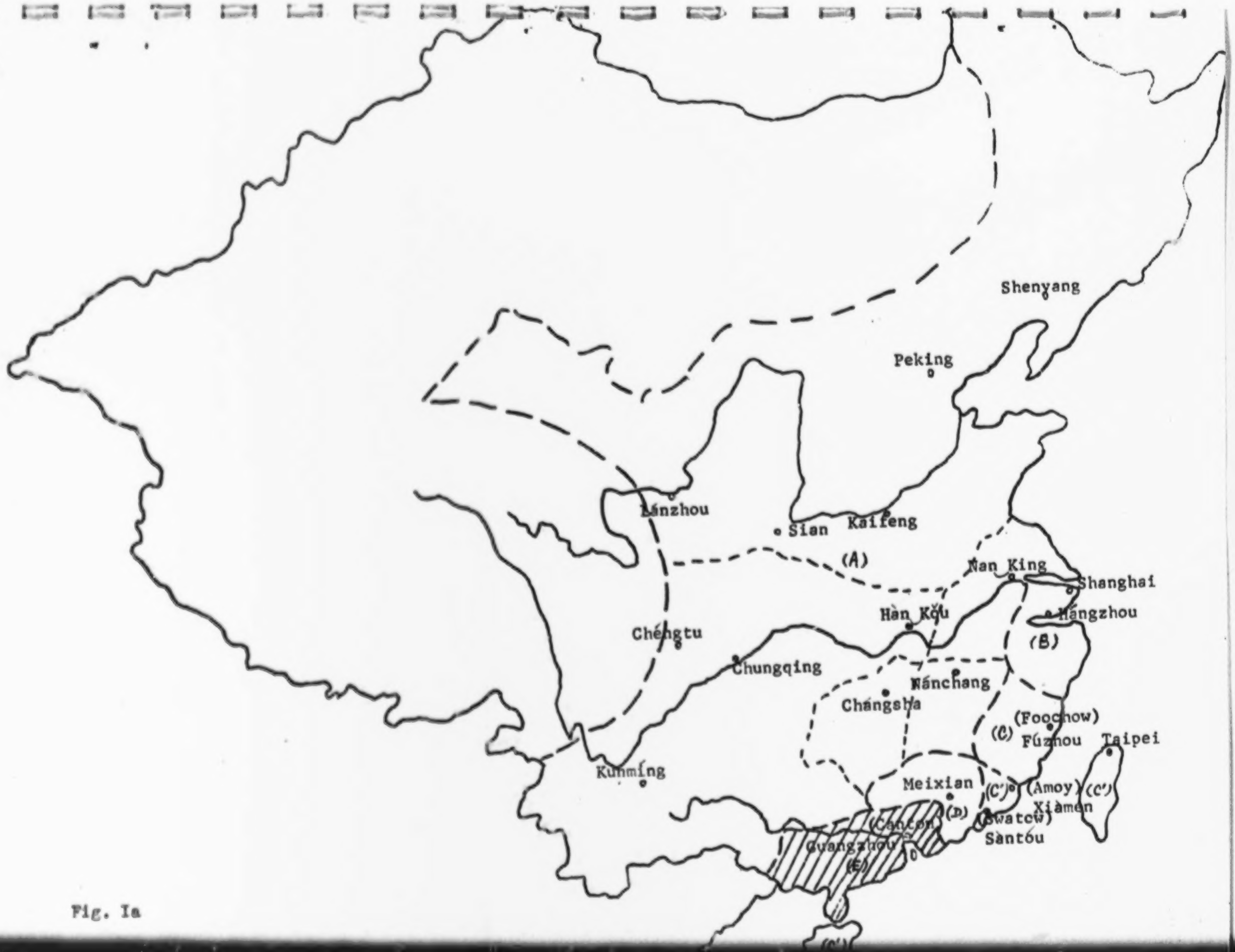

Shenyang
Peking
Lánzhou
Sian
Kaifeng
(A)
Nan King
Shanghai
Hán Kǒu
Hángzhou
(B)
Chéngtu
Chungqing
Nánchang
Chángsha
(Foochow)
(C)
Fúzhou
Taipei
Kunming
Meixian
(C')
(Amoy)
(C')
Xiàmen
(D)
(Swatow)
Sàntóu
Canton
Guangzhou
(E)
(C')

Fig. Ia

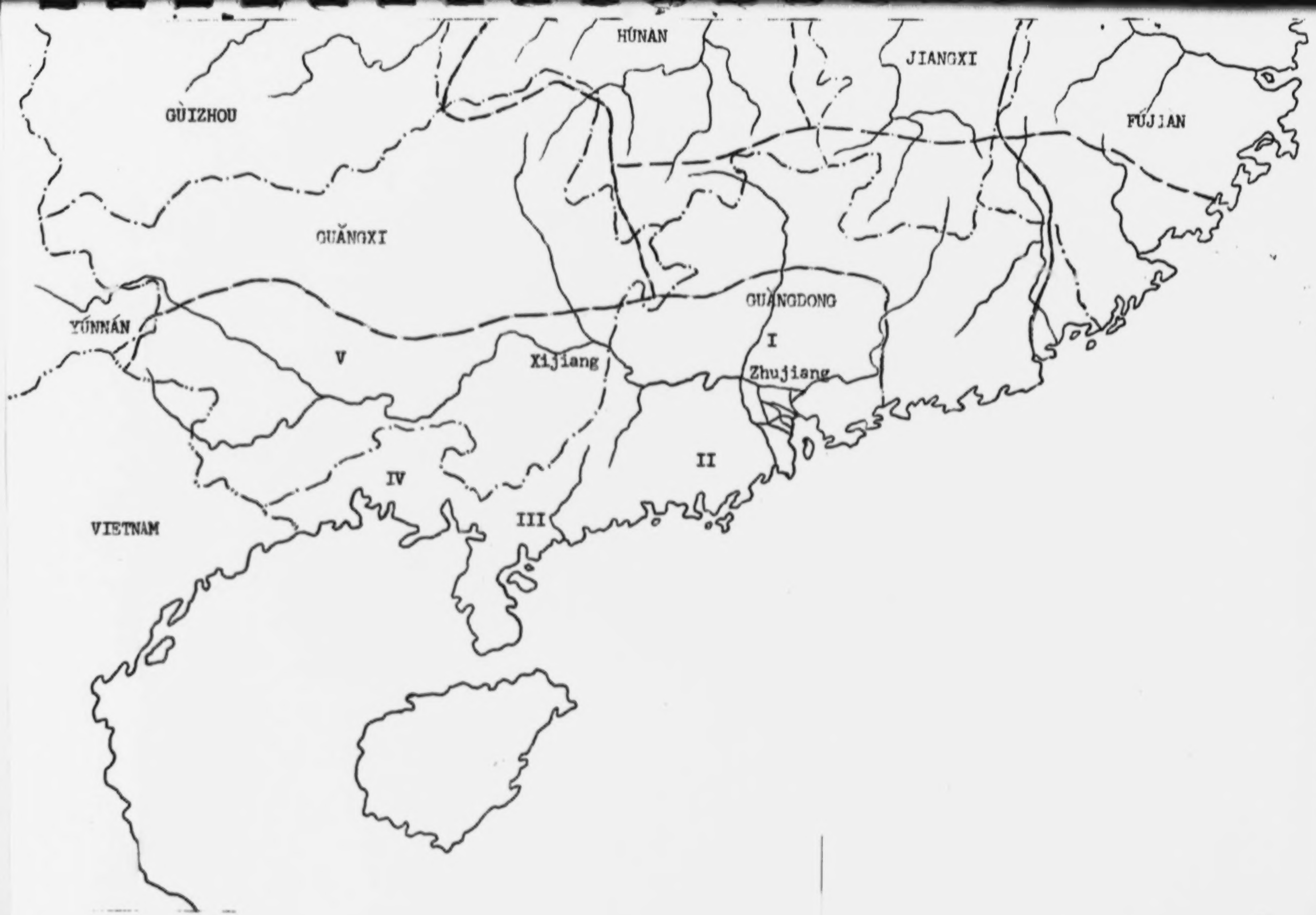

HUNAN
JIANGXI
GUIZHOU
FUJIAN
GUĂNGXI
GUĂNGDONG
YÚNNÁN
I
V
Xijiang
Zhujiang
II
IV
III
VIETNAM

Fig. Ib

C5

Bǎisè
Wúzhou
Nánníng
(Yungníng)
Róngxiàn
Yùlín
Nánhǎi
Guǎngzhou
(Canton)
Panyú
Lúngzhou
Shùngdé
Bóbái
Xinhùi
Zhongshan
Kowloon
Hongkong
Kaipíng
Táishan
Qinzhou
Enpíng
Gaozhou
(Màomíng)
VIETNAM
Liánzhou
(Hépǔ)
Zhanjiang
SOUTH CHINA
SEA
Léizhou
GULF OF TONKIN
HǍINÁN

Fig. Ic

Chapter I. Introduction

1.1 The Cantonese dialect. The term 'Cantonese', in its broader sense, refers to a group of similar dialects spoken in southern China, in the southern half of the Guǎngxi (Kwangsi)[a] province and the western half of the Guǎngdong (Kwangtung)[b] province, and often referred to as the Yùe[c] dialects.[1] Settlements of speakers of this group are found in Southeast Asia, in the Hawaiian Islands, and in North, Central, and South America. Cantonese claims a large number of overseas Chinese speakers.

The dialect group called Cantonese is characterized by:

(1) the preservation of the final consonants -m, -n, -ŋ, the yáng[d] final series, and -p, -t, -k, the rù[e] final series; the latter is so called because its members co-occur with the rù (or entering) tones;

(2) the presence of nine or more tonal contrasts for isolated syllables at the systematic phonetic level; and

(3) the distinction between long and short vowels interplaying with the distinction among several tones.

Within this dialect group we can again distinguish five subgroups according to dialectal differences and areas of distribution.[2] Of these five subgroups, the Yùe hǎi[f] subgroup is the most important because it includes the lingua franca of the area, the so-called standard Cantonese, among its members. Standard Cantonese is the dialect of Canton City and its closest variety, spoken in Kowloon and Hongkong. The term 'Cantonese', in its narrower sense, refers to standard Cantonese; it is in this sense that Cantonese will be used throughout this paper.

1.2 Theoretical background of analysis. The present study of the Cantonese phonological system is meant to be a case study of the phonological

component of a generative grammar. A generative grammar is a system of sound-meaning correlations, in the form of a set of rules which assign structural descriptions to the sentences of a particular language and thereby express a fundamental aspect of the speaker's competence. This system may be represented graphically as follows:

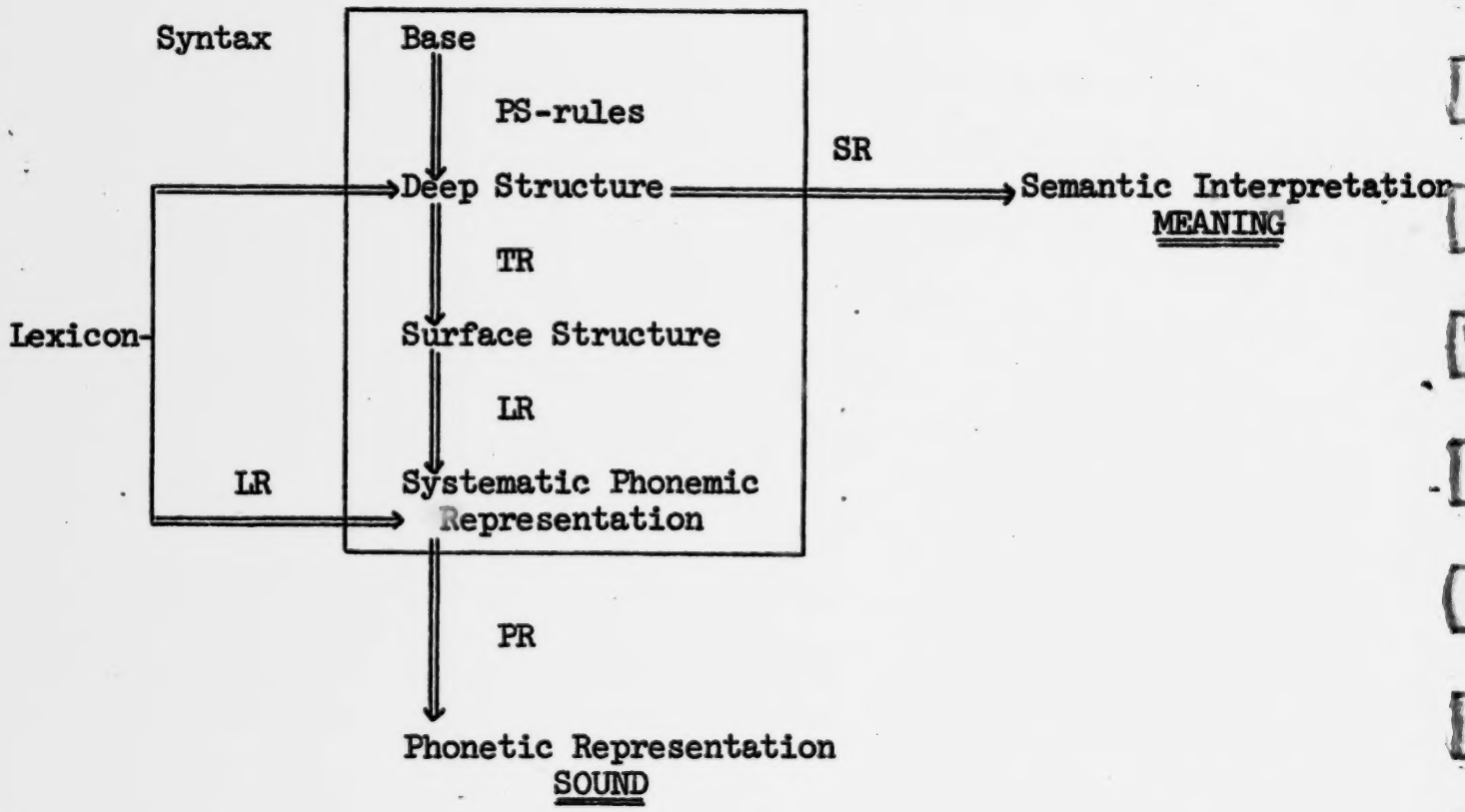


Figure II. A model of sound-meaning correlations.

It is taken for granted that the reader is familiar with Chomsky's Aspects of the theory of syntax, so that we can go directly into the phonological component which is our primary concern. The phonological component of a generative grammar corresponds to the lower part of the graph where the phonetic representation of an utterance is shown to be connected with syntax on the one hand and the lexicon on the other. Roughly speaking,

three things are involved in the phonological component which yield the ultimate phonetic representation of an utterance: (1) the structural description or phrase marker in the form of labeled bracketing; (2) the systematic phonemic representation of each lexical entry introduced by the lexical rule; and (3) the phonological rules.

Both the lexical entries and the phonological rules are expressed through the use of distinctive features. The function of distinctive features is to analyze each sound into its essential features, which are responsible for sameness or difference and for the interactions among sounds. It is the task of a universal grammarian to provide us with a repertory of distinctive features with which we can describe any language in the world.

Phonological rules may be either syntactically or phonologically motivated. Hence it is rather meaningless to speak of a sharp demarcation between the syntactic and the phonological component of a generative grammar, as may be suggested by the graph in Figure II. The structure of a sentence and the grammatical categorization of each of its lexical entries have definite conditioning effects on its pronunciation. The interaction of sounds or, more precisely, the interaction of features also contributes to the pronunciation. In the Indo-European languages, as well as in other language families, we have numerous examples of sandhi, i.e. phonologically motivated sound changes.

The interaction of features may sometimes be attributed to the inherent redundancy built into the universal system by the very definitions of the phonological distinctive features. This kind of redundancy is generally understood and need not be expressed outside the universal grammar. Hence, we may say that this type of feature interaction is a natural consequence

of the overlapping of domains of phonological features. Another type of feature interaction which is accidental and differs from language to language generally results from the morpheme structures peculiar to each language. The task of phonological rules is to make explicit the latter type of feature interaction, which includes sequential constraint among features and interaction between segmental and prosodic features.

As opposed to segmental features, prosodic features indicate phonological properties that range over the entire syllable or morpheme. Hence they are also called suprasegmental features. Prosodic features interact with segmental features at various places in the syllable or morpheme. In the case of pitch or tonal features, there is a pretty good physical correlation, in the sense that tonal features range over the entire voiced portion of the syllable. However, this is not true of all prosodic features. For example, a diacritic feature like [± Romance] or [± colloquial] does not have any physical correlation. It is introduced solely to simplify the system by capturing some regularities governing the behavior of a class of morphemes.

Chapter II. The sounds of standard Cantonese

2.1 Major class features of Cantonese segmental sounds. Since an identical set of sounds may alternate differently in different phonological systems, empirical considerations are of primary importance in the presentation of major classes or the basic configuration of the system. One key to major classification lies in the classification of the nasal consonants. [n] alternates with [l] in Cantonese and some Húnán dialects, but with [d] in Swahili; while in Taiwanese, a Min dialect, [l] alternates with [d]. In Figure III we present the major classes of the three different systems:

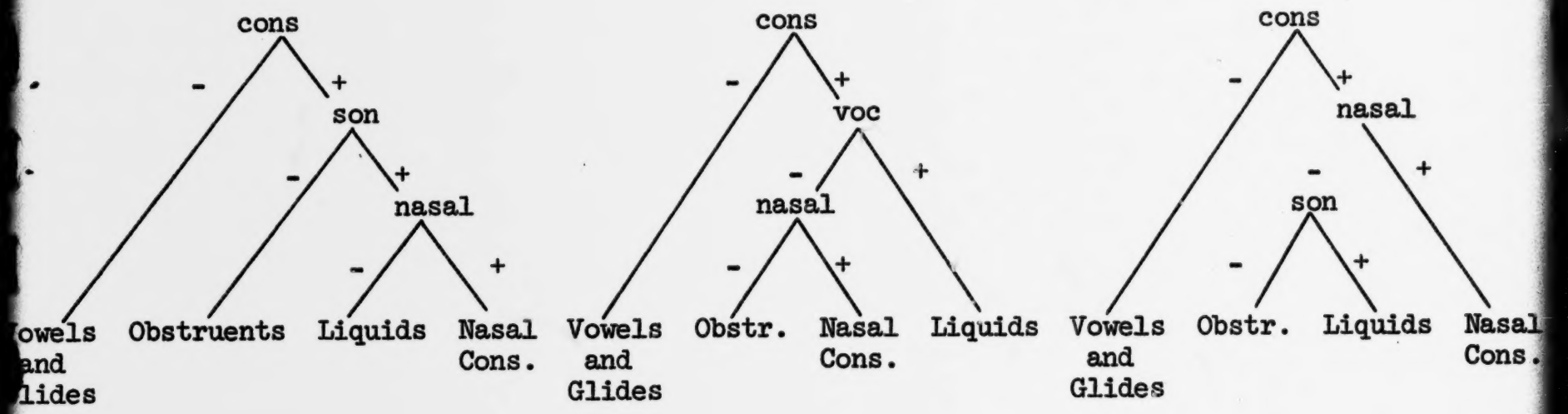


Figure III. Natural classification of the nasal consonants.

In each system three features are involved. They are [± consonantal], [± nasal], and [± sonorant] or [± vocalic]. It is the consideration of [n] and its natural class that leads us to the above selections. For Cantonese, however, [± vocalic]--in Swahili, [± sonorant]--may be substituted for the feature [± nasal]. In doing this we are using the three major class

features, i.e. [± consonantal], [± sonorant] and [± vocalic], proposed by Chomsky and Halle which, unfortunately, fail to adequately account for the configuration of Taiwanese. Since we need the feature [± vocalic] to further separate the vowels from the glides in our study of Cantonese, we shall favor the use of [± nasal]. Thus the major classes of Cantonese form the following configuration:

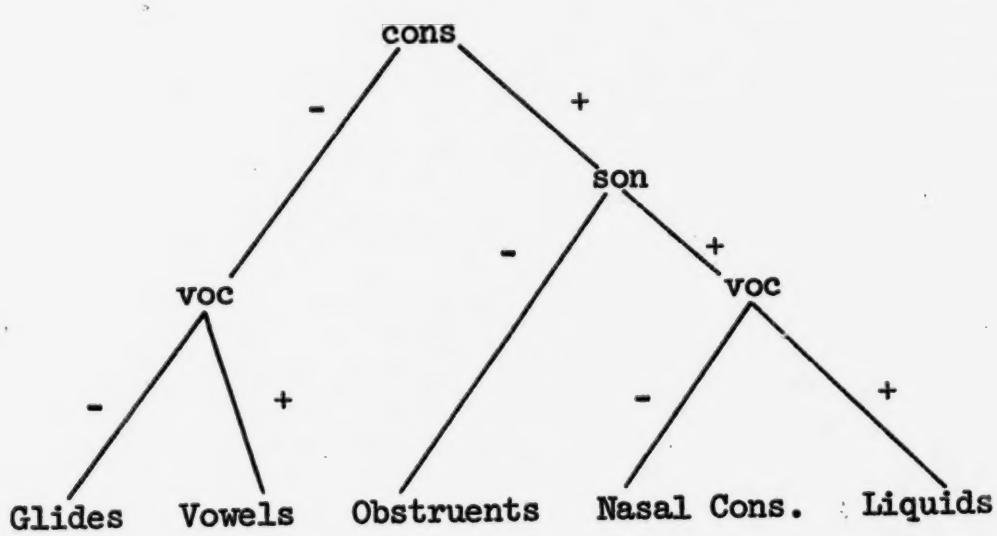


Figure IV. Major classes of Cantonese sounds.

2.2 The segmental system of Cantonese. To assign a unique representation to each segmental sound of Cantonese, we use the following features suggested by Chomsky and Halle in The sound pattern of English:[3]

A. Major class features

consonantal: obstruction of air in the vocal tract

sonorant: free passage of air through either the vocal tract or the nasal tract

vocalic: free passage of air through the vocal tract only, produced with a periodic excitation accompanied by open oral cavity

B. Cavity features

grave-acute: grave sounds are articulated with a primary narrowing located at the periphery of the oral cavity; i.e. at the lips or in the velar or pharyngeal region

diffuse-nondiffuse: diffuse sounds are produced with a narrowing which equals or exceeds that of a constriction and is located in the front part of the vocal tract

compact-noncompact: this feature is restricted to vowels. Compact vowels are produced with a forward and downward flanged oral cavity which contains no constriction or narrowing of high degree

flat-nonflat: flat sounds are produced with a secondary narrowing at the periphery of the oral cavity often referred to as rounding

C. Manner of articulation features

continuant-discontinuous: continuant sounds are characterized by the absence of abrupt transition between sound and silence

tense: tense sounds are produced with greater deformation of the vocal tract away from its rest position and emit a higher amount of energy in conjunction with a longer duration in time. In stop consonants, tenseness is manifested by a greater strength of the explosion.

D. Source feature

strident: complex impediment, rough-edged supplementary obstruction creating edge effects at the point of articulation

The above-mentioned features give us a Cantonese phonological system consisting of eleven obstruents or true consonants, three nasal consonants, one liquid, eight vowels, two glides, and the glottal stop (here presented

in Figure V). Obstruents and vowels and their variants will be treated in detail in Chapter III.

2.3 Redundancy rules. With these rules we have filled in certain holes in the matrix given in Figure Va. (Features generated by this set of redundancy rules are shown in Figure Vb.)

(1) $[-\text{son}] \longrightarrow [-\text{voc}]$

(2) $[+\text{voc}] \longrightarrow [+\text{son}]$

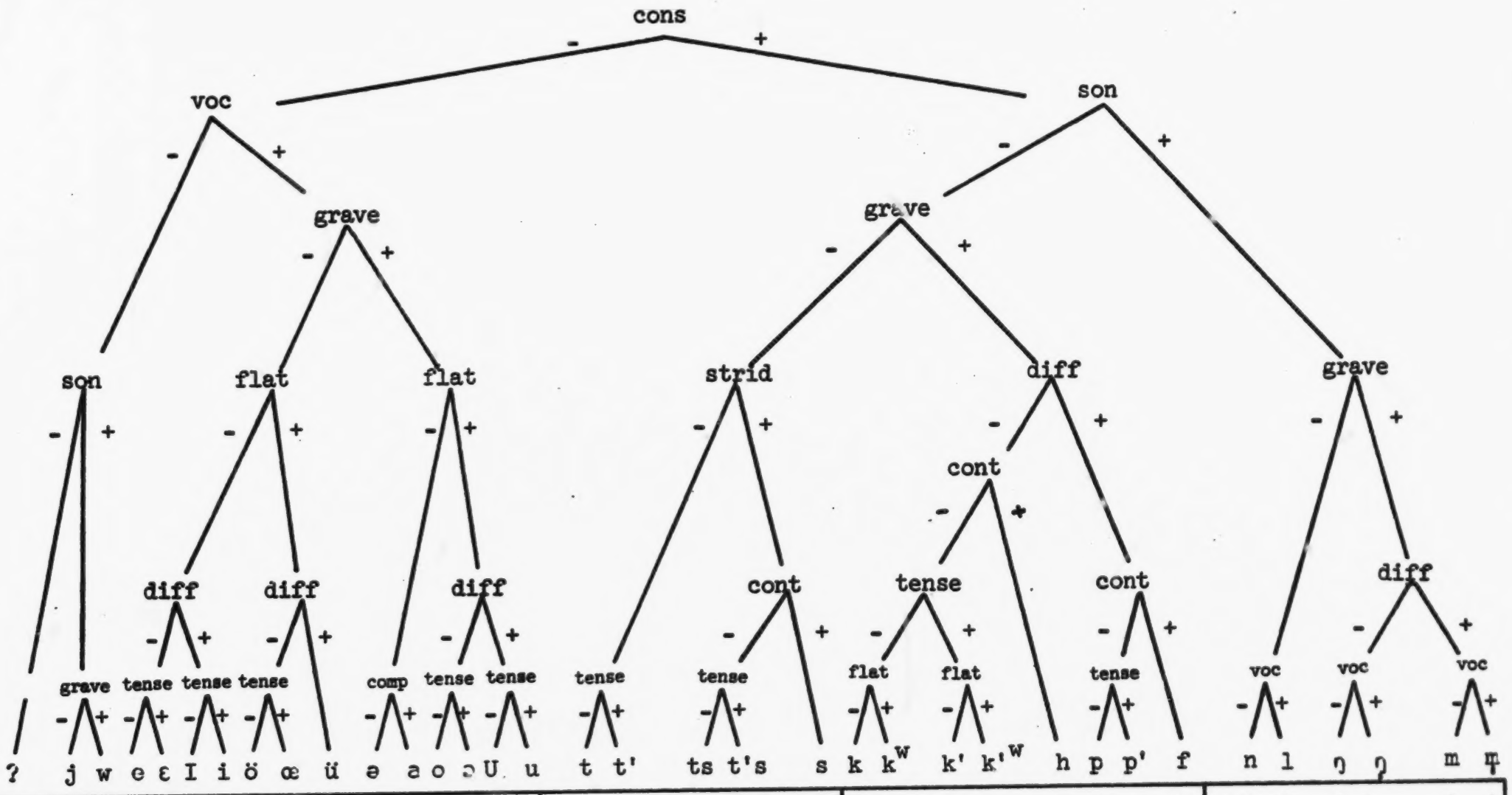
(3) $\left\{ \begin{matrix} [+\text{son}] \\ [+\text{voc}] \end{matrix} \right\} \longrightarrow [+\text{cont}]$

(4) $\begin{bmatrix} -\text{cons} \\ -\text{voc} \\ -\text{son} \end{bmatrix} \longrightarrow \begin{bmatrix} +\text{grave} \\ -\text{flat} \end{bmatrix}$

(5) $[\alpha\text{son}] \longrightarrow \begin{bmatrix} \alpha\text{diff} \\ \alpha\text{cont} \\ -\alpha\text{tense} \end{bmatrix} / \begin{bmatrix} -\text{cons} \\ -\text{voc} \\ ____ \end{bmatrix}$

(6) $[\alpha\text{comp}] \longrightarrow [\alpha\text{tense}] / \begin{bmatrix} -\text{cons} \\ +\text{voc} \\ +\text{grave} \\ -\text{flat} \\ ____ \end{bmatrix}$

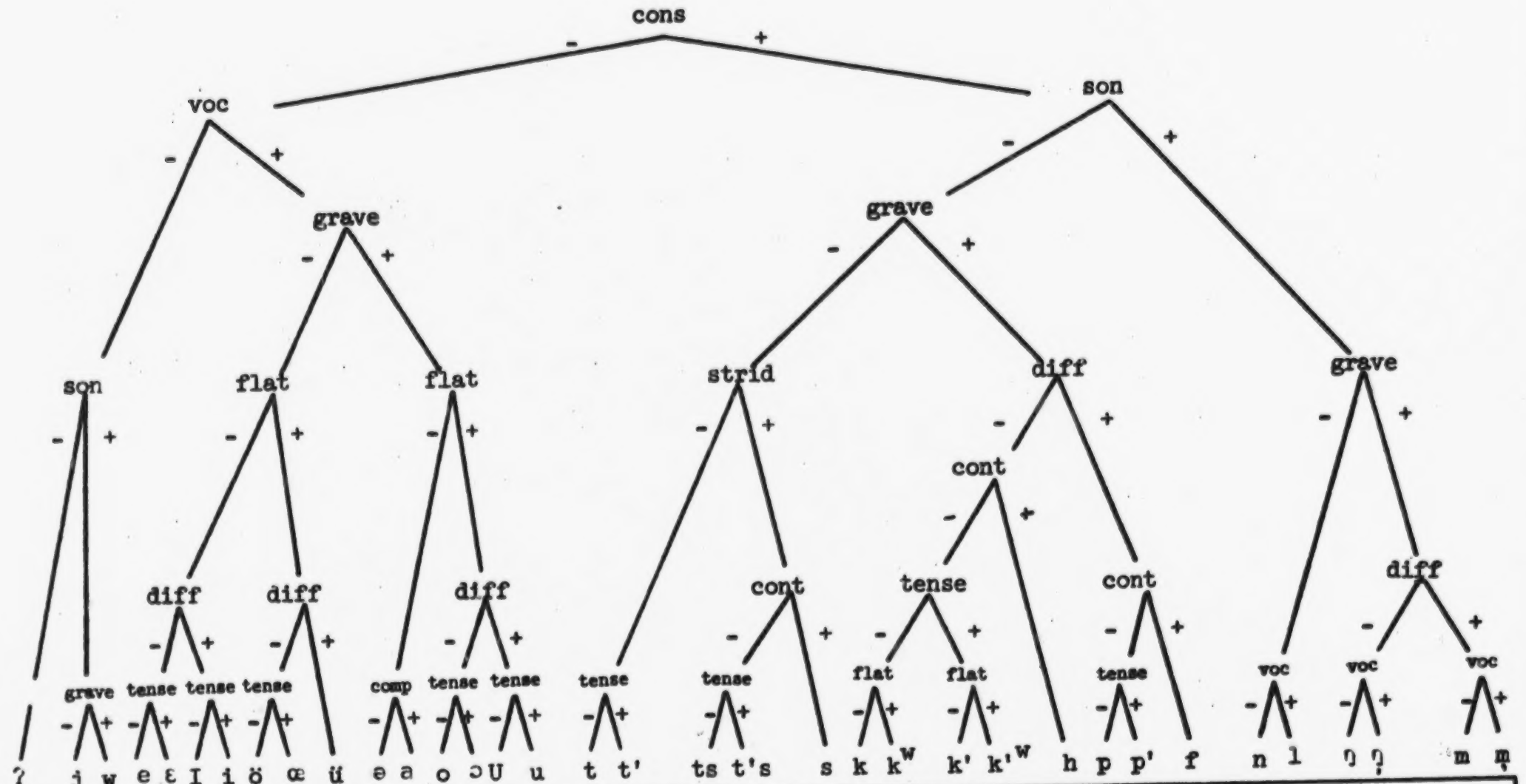
(7) $[\alpha\text{cons}] \longrightarrow [-\alpha\text{tense}] \begin{bmatrix} +\text{cont} \\ ____ \end{bmatrix}$



	?	j	w	e	ɛ	I	i	ö	œ	ü	ə	ɑ	o	ɔ	U	u	t	t'	ts	t's	s	k	k^w	k'	k'^w	h	p	p'	f	n	l	ŋ	ʔ	m	ɱ
consonantal	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+
sonorant	-	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	-	-	-	-	-	-	-	-	-	-	-	-	-	-	+	+	+	+
vocalic	-	-	-	+	+	+	+	+	+	+	+	+	+	+	+	+	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	+	-	+
grave	+	-	+	-	-	-	-	-	-	-	+	+	+	+	+	+	-	-	-	-	-	+	+	+	+	+	+	+	+	+	-	-	+	+
flat	-	-	+	-	-	-	-	+	+	+	-	-	+	+	+	+	-	-	-	-	-	-	+	-	+	-	-	-	-	-	-	-	-	-
diffuse	-	+	+	-	-	+	+	-	-	+	-	-	-	-	+	+	+	+	+	+	+	-	-	-	-	-	-	+	+	+	+	-	+	+
strident	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	+	+	+	-	-	-	-	-	-	-	-	-	-	-	-	-
compact	-	-	-	-	-	-	-	-	-	-	-	-	+	-	+	-																		
continuant	-	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	-	-	-	-	+	-	-	-	-	+	+	-	-	+	+	+	+	+
tense	+	-	-	-	+	-	+	+	-	+	-	+	-	+	-	+	-	+	-	+	-	-	-	+	+	-	-	+	-	-	-	-	-	-

Figure Va. The segmental system of Cantonese.

Cl4a


cons
voc
son
grave
grave
son
flat
flat
strid
diff
grave
diff
diff
diff
cont
cont
tense
cont
diff
grave
tense
tense
comp
tense
tense
tense
tense
flat
flat
tense
voc
voc
voc
? j w e ɛ I i ö œ ü ə a o ɔ U u t t' ts t's s k k^w k' k'^w h p p' f n l ŋ ʔ m m̥

| | | |

(8) $\begin{bmatrix} -\text{cons} \\ +\text{voc} \\ +\text{grave} \\ +\text{flat} \\ -\text{diff} \\ +\text{tense} \end{bmatrix} \longrightarrow [+\text{comp}]$

(9) [ ] $\longrightarrow$ [-comp]

(10) [αgrave] $\longrightarrow$ [αflat] / $\begin{bmatrix} -\text{cons} \\ +\text{son} \\ -\text{voc} \\ ____ \end{bmatrix}$

(11) [ ] $\longrightarrow$ [-flat]

(12) [αcons] $\longrightarrow$ [-αdiff]

(13) [ ] $\longrightarrow$ [-strid]

(14) [+strid] $\longrightarrow$ [-grave]

2.4 <u>The tonal system of Cantonese</u>. To assign a unique representation to each Cantonese tone, we use the tone features suggested by William S-Y. Wang (1967:97-9), with the following set of redundancy conventions:

(1) [-contour] $\longrightarrow$ $\begin{bmatrix} -\text{rising} \\ -\text{falling} \\ -\text{convex} \end{bmatrix}$

(2) $\left\{ \begin{matrix} [+\text{high}] \\ [-\text{central}] \end{matrix} \right\} \longrightarrow [-\text{mid}]$

(3) [+central] $\longrightarrow$ [-contour]

(4) $\left\{\begin{matrix}[+\text{rising}]\\[+\text{falling}]\end{matrix}\right\} \longrightarrow [+\text{contour}]$

(5) $[+\text{contour}] \longrightarrow \begin{bmatrix}-\text{central}\\-\text{mid}\end{bmatrix}$

(6) $\left\{\begin{matrix}[-\text{rising}]\\[-\text{falling}]\end{matrix}\right\} \longrightarrow [-\text{convex}]$

We find from the Catonese data that ˥ 55 alternates with ˧˥53; ˧˥ 35 with ˨˩ 21, ˨˧ 22, ˧ 33, ˧˩ 23; and ˦ 4 with ˩ 2 and ˧ 3. With the present set of tone features, and the consideration of natural classes, we are in a difficult position to classify the Cantonese tones, since the latter two alternations each involves several tones and the change of at least two features. Thus, tones belonging to the same natural class may be far apart in the configuration shown in Figure VI.

2.5 Redundancy rules.

(1) $[+\text{high}] \longrightarrow \begin{bmatrix}-\text{central}\\-\text{mid}\end{bmatrix}$

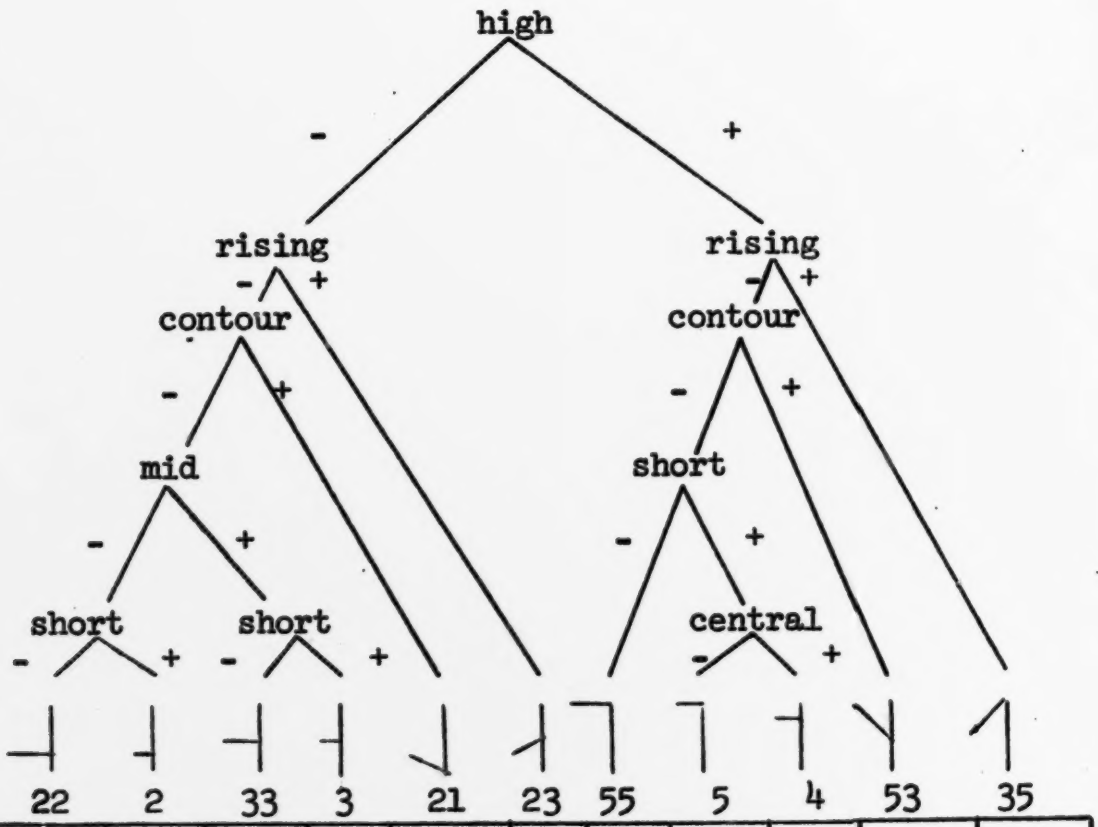
(2) $[+\text{mid}] \longrightarrow [+\text{central}]$

(3) $[-\text{central}] \longrightarrow [-\text{mid}]$

(4) $\begin{bmatrix}-\text{short}\\+\text{rising}\end{bmatrix} \longrightarrow [+\text{contour}]$

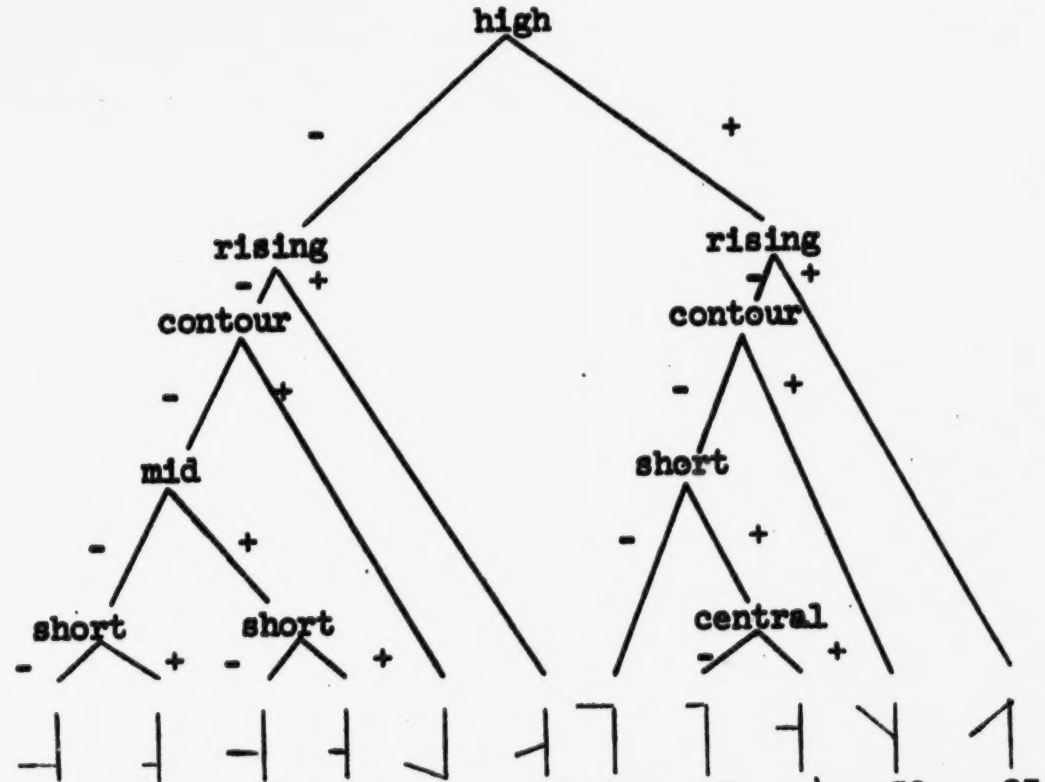
(5) $[+\text{contour}] \longrightarrow [-\text{short}]$

Due to the interaction of pitch, duration, and final consonants, it is not generally necessary to mark the feature [± short] on the systematic phonemic level. This feature is introduced here only to facilitate further descriptions.



	22	2	33	3	21	23	55	5	4	53	35
±high	-	-	-	-	-	-	+	+	+	+	+
±rising	-	-	-	-	-	+	-	-	-	-	+
±contour	-	-	-	-	+	+	-	-	-	+	+
±central	+	+	+	+	-	-	-	-	+	-	-
±mid	-	-	+	+	-	-	-	-	-	-	-
±short	-	+	-	+	-	-	-	+	+	-	-

Figure VIa. The tonal system of Cantonese.



	22	2	33	3	21	23	55	5	4	53	35
±high											
±rising											
±contour						+					+
±central	+	+	+	+	-	-	-			-	-
±mid					-	-	-	-	-	-	-
±short					-	-				-	-

Figure VIb. The tonal system of Cantonese: features generated by redundancy rules.

Chapter III. Problems in the treatment of Cantonese phonology

3.1 Some deviations from standard Cantonese. Even though one has to idealize away from idiolectic and idiosyncratic elements in the speech of various members of the speech community when making linguistic descriptions, there are nonetheless three variant features of Cantonese pronunciation that are worth mentioning (see Lǐ 1937):

(1) the absence of a distinction between initial l and n

(2) the confusion of the zero initial and initial ŋ

(3) the use of the apical vowel

3.1.1 The absence of a distinction between initial l and n. Historically speaking, Cantonese initial n results from the merger of two Ancient Chinese initials, ní[g] and niáng,[h] while Cantonese initial l comes from the Ancient Chinese initial lai.[i] The distinction between n and l is evidenced by Hakka,[j] Min,[k] and Wu[l] dialects of Chinese, as well as by Mandarin. One would expect that in Cantonese also this distinction would be kept for the two sets of words with different origins. This is not the case, however, in the speech of some Cantonese speakers, especially those from the vicinity of Nán-hǎi.[m] The two sets of words are pronounced either with the same initial, l or n, or with l and n in free variation with each other.

3.1.2 The confusion of the zero initial and initial ŋ. The rime dictionaries indicate that words belonging to the Yǐng[n] group share an initial distinct from that of the Yi[o] group in Ancient Chinese. Though there has been merging through the years, this distinction is still evidenced by some Hakka, Min, and Wu dialects. In Cantonese, however, the situation is rather chaotic. The two initials neither merge nor remain

distinct. The two sets of words are pronounced in all the possible ways: (a) some speakers keep the distinction; (b) some pronounce both sets with initial ŋ; (c) some drop initial ŋ from the system and pronounce both sets with the zero initial; and (d) some may even pronounce words of the Yíng group with initial ŋ and words of the Yí group with the zero initial, just the reverse of what is considered standard.

3.1.3 The use of the apical vowel. In Cantonese, the sounds which are transcribed /ts, t's, s/ in Figure V are ordinarily pronounced as alveopalatal continuants: [tɕ, t'ɕ, ɕ]. However, some educated and older Cantonese speakers tend instead to use the dental counterparts of these sounds when they precede the apical vowel /i/; /i/ in turn becomes [ɿ], thus yielding the sounds [tsɿ, t'sɿ, sɿ]. This pronunciation was once prestigious, but is now little used by the younger generation. The lingering effect of this phenomenon is reflected in the replacement of the palatal series by the dental series in the speech of some speakers. Since the two series never oppose each other within the same system, we shall for ease of transcription choose as norm the symbols for the dentals and expect no confusion to arise from our preference.

3.2 Initials. Here the terms initial, medial, final, and tone are used as in the traditional classification of Chinese sounds in rime dictionaries. However, in accordance with our segmental analysis, a final should be represented as two segments, consisting of the main vowel optionally followed by a glide, a nasal, or a stop consonant. Since our primary concern here is the treatment of Cantonese sounds, we shall leave the discussion of syllable structure in terms of segmental and tonal characterization until Chapter IV.

3.2.1 The zero initial. In Ancient Chinese, the zero initial belongs to the Yú^{p} group. In Cantonese, what we call the zero initial results from

the merging of the Yú group (with initial *ø or *w), the Yíng group (with initial *ʔ), the Yún[q] group (with initial *j), and the Wéi[r] group (with initial *ŋ). Here we choose to keep distinct the zero initial and initial ŋ for Cantonese, though a further step of merging has already taken place for Mandarin. The zero initial of Cantonese is realized phonetically as a semivowel (j, w) or as the glottal stop. Though in the speech of some speakers the glottal stop is sometimes lost, we shall still mark it in our transcription. On the systematic phonemic level, however, it is not necessary to represent the initial semivowels since they can be predicted from the following medial or main vowel. In view of this, the zero initial is justified on the systematic phonemic level. This underlying 'zero' may be the explanation for the gradual disappearance of initial ŋ, ʔ, j, and w in Cantonese, a process almost completed by Mandarin.

Starting with a zero initial in the systematic phonemic representation, the following instructions generate the respective initials. If a syllable starts with medial -u-, i.e. if a syllable lacks an initial segment marked [+cons], and if -u- is not followed by the vowel /i/, initial w replaces the medial -u-. A medial -i- not preceded by an initial bearing the [+cons] feature and not followed by the vowel /u/ is replaced by initial j. If /u/ is the main vowel of a syllable not preceded by a [+cons] initial, initial w is introduced. If /i/, or /ü/, is the main vowel and is not preceded by a [+cons] initial, initial j is introduced. When u combines with i in the forms ui and iu, the first element is treated as the main vowel rather than as a medial. The initial semivowels w and j are introduced accordingly.

Finals consisting of other main vowels, when not preceded by any of the [+cons] initials, take as their initials the glottal stop or ŋ: those

with yin tones take the glottal stop; those with yáng tones take ŋ. The zero initial, in this sense, no longer exists on the systematic phonetic level, though the recognition of its existence on the systematic phonemic level may have greatly simplified our system.

We can summarize the above operations in the following ordered rules:

(1) $\emptyset \rightarrow \begin{bmatrix} -\text{cons} \\ -\text{voc} \\ +\text{son} \\ \alpha\text{grave} \end{bmatrix} / \# ____ \begin{bmatrix} -\text{cons} \\ +\text{voc} \\ \alpha\text{grave} \\ +\text{diff} \end{bmatrix} \left(\begin{bmatrix} -\text{cons} \\ \pm\text{voc} \end{bmatrix} \right)$

(2) $\begin{bmatrix} -\text{cons} \\ +\text{voc} \\ +\text{diff} \\ \alpha\text{grave} \end{bmatrix} \rightarrow \emptyset / \# \begin{bmatrix} -\text{cons} \\ -\text{voc} \\ +\text{son} \\ \alpha\text{grave} \end{bmatrix} ____ \begin{bmatrix} -\text{cons} \\ +\text{voc} \end{bmatrix}$

(3) $\emptyset \rightarrow \begin{bmatrix} +\text{cons} \\ +\text{son} \\ +\text{grave} \\ -\text{diff} \\ -\text{voc} \end{bmatrix} / \# ____ \begin{bmatrix} -\text{cons} \\ +\text{voc} \end{bmatrix}$ with yáng tones

(4) $\emptyset \rightarrow \begin{bmatrix} -\text{cons} \\ -\text{voc} \\ -\text{son} \end{bmatrix}$ with yin tones

The zero initial which functions like an initial consonant is, in fact, non-consonantal in three of its four possible realizations. Hence, it should be distinguished from the stop consonants and the nasal consonants. Here we introduce a diacritic feature [± consonantal'] which should be regarded as an equivalent of [± consonantal] with respect to its function in

sequential constraints, but in feature specification of lexical entries this feature has the following implications:

(a) [u cons'] ⟶ [-cons']

(b) [+cons'] ⟶ $\begin{bmatrix}-\text{cons}\\-\text{voc}\\+\text{son}\\\alpha\text{grave}\end{bmatrix}$ / # ___ $\begin{bmatrix}-\text{cons}\\+\text{voc}\\\alpha\text{grave}\\+\text{diff}\end{bmatrix}$ $\left(\begin{bmatrix}-\text{cons}\\+\text{voc}\end{bmatrix}\right)$

(c) [+cons'] ⟶ $\begin{bmatrix}+\text{cons}\\+\text{son}\\+\text{grave}\\-\text{diff}\\-\text{voc}\end{bmatrix}$ / # ___ $\begin{bmatrix}-\text{cons}\\+\text{voc}\end{bmatrix}$ with yáng tones

(d) [+cons'] ⟶ $\begin{bmatrix}-\text{cons}\\-\text{voc}\\-\text{son}\end{bmatrix}$ elsewhere

(b), (c), and (d) are in fact modifications of rules (1), (3), and (4). In referring to rules (1)-(4) of this section, we mean the modified version of these rules. From now on, an initial consonantal segment must be marked either [+cons] or [+cons']. A nonconsonantal segment must be [-cons] and [+voc].

3.2.2 <u>The initials n, l, ŋ, and ʔ (∅)</u>. In section 3.1 we have already touched upon the absence of a distinction between initial n and l and between ŋ and ʔ(∅) in the speech of some Cantonese speakers. Linguists who take into consideration the historical origin of such initials often insist upon keeping them distinct.[4,5] They simply explain away the

chaotic situation mentioned above as the coexistence of standard and deviant pronunciations. A linguist's job, however, is to record the actual situation and explain whatever changes he finds, instead of assuming no change at all and explaining any change as deviant. As the case of the zero initial may indicate, alternation and variation of sounds may well be symptoms of a sound change in progress.

So far, the arbitrariness of linguists and lexicographers has rarely been questioned. This may be because words with initial n are few in number. The list given in footnote s almost exhausts the inventory except for a few homophones of the words already given. A literate person can easily memorize the list. This may also explain the fact that nonnative speakers of Cantonese can do better in a test of l and n distinctions than some native Cantonese speakers.

To assume a distinction between initial ŋ and ʔ(ø) is arbitrary: not only has empirical evidence been ignored; the historical argument is also unsound. In view of the development of Ancient Chinese into Mandarin, it is not unlikely that, in Cantonese also, initial *ŋ can be dropped by some members of the Yí group, depending on its environment. Comparing qié-yùn[t] initials with corresponding ones in modern Cantonese pronunciation, we can find many instances where initial *ŋ and *ʔ (ø) merge with the zero initial. For example (Wáng 1957:127):

*ŋiu[u] → ü ← *ʔiu[v]

*ŋan[w] → on ← *ŋan[x]

*ŋau[y] → ou ← *ŋau[z]

This is to say, the merging of the Yí and Yǐng groups actually takes place for Cantonese as well as for Mandarin. There is absolutely no reason why the qié-yùn classification should be strictly followed in the classification should be strictly followed in the classification of sounds in modern

Chinese dialects. This type of classificational problem is, however, not of central interest in our framework.

3.2.3 Labialized stops vs. the medial -u-. This problem centers around the status of two labialized velars: k^{W} and k'^{W}. This is by no means a problem unique to Cantonese. Parallels are found in the analysis of Mandarin and Amoy. We find it unnecessary to set up additional labialized velars in our system. The reasons are:

(1) kuX is not distinct from k^{W}uX, nor is k'ux distinct from k'^{W}uX (X may be null or any permissible final),

(2) the kaikǒu hu[ac] medial quality can be represented by the medial -u- or the main vowel /u/ without being attributed to the initial stop, while

(3) the secondary rounding feature of the initial k or k' can be predicted by the following segment, and

(4) secondary rounding is just not a common feature for Chinese initials, so it is best to consider it a result of assimilation to the following rounded vowel.

If we were classifying the Cantonese sounds in terms of initials and finals, we might have to worry about the fifteen extra kaikou hu finals (-ua, -uəi, -uai, -uo, -uən, -uan, -uəŋ, -uaŋ, -uoŋ, -uiŋ, -uət, -uat, -uak, -uok, and -uik) to be added to the final inventory as a result of excluding two initials, k^{W} and k'^{W}. But within the present theoretical framework, simplicity is not measured by the number of distinctive segments, but by the overall patterning in terms of distinctive features.

3.3 Medials. There are only two medials in Cantonese: -u- and -i-. Medial -u- occurs only after initial k, k', or ∅. The semivowel w is introduced in the place of ∅ when followed by -u-. Medial -i- always

occurs with ∅ which is realized as the semivowel j. Eventually both medials are dropped. This operation can be represented by the following two ordered rules:

(1) $\emptyset \rightarrow \begin{bmatrix} -\text{cons} \\ -\text{voc} \\ +\text{son} \\ \alpha\text{grave} \end{bmatrix} \; / \; \# __ \begin{bmatrix} -\text{cons} \\ +\text{voc} \\ \alpha\text{grave} \\ -\text{diff} \\ -\text{grave} \end{bmatrix} \left(\begin{bmatrix} -\text{cons} \\ +\text{voc} \end{bmatrix} \right)$

(2) $\begin{bmatrix} -\text{cons} \\ +\text{voc} \\ +\text{diff} \\ \alpha\text{grave} \end{bmatrix} \rightarrow \emptyset \; / \; \# \begin{bmatrix} -\text{cons} \\ -\text{voc} \\ +\text{son} \\ \alpha\text{grave} \end{bmatrix} __ \begin{bmatrix} -\text{cons} \\ +\text{voc} \end{bmatrix}$

When -u- is preceded by a velar stop, the following operation takes place:

(5) $[__] \rightarrow [+\text{flat}] \; / \; \# \begin{bmatrix} +\text{cons} \\ -\text{son} \\ -\text{grave} \\ -\text{diff} \\ -\text{cont} \\ ____ \end{bmatrix} \begin{bmatrix} -\text{cons} \\ +\text{voc} \\ +\text{grave} \\ +\text{flat} \\ +\text{diff} \end{bmatrix} \left(\begin{bmatrix} -\text{cons} \\ +\text{voc} \end{bmatrix} \right)$

(6) $\begin{bmatrix} -\text{cons} \\ +\text{voc} \\ +\text{grave} \\ +\text{flat} \\ +\text{diff} \end{bmatrix} \rightarrow \emptyset \; / \; \# \begin{bmatrix} +\text{cons} \\ -\text{son} \\ +\text{grave} \\ -\text{diff} \\ -\text{cont} \\ +\text{flat} \end{bmatrix} __ \begin{bmatrix} -\text{cons} \\ +\text{voc} \end{bmatrix}$

In other words, k or k' takes on the secondary rounding feature and the conditioning factor, i.e. medial -u-, eventually drops.

If we understand a medial as a segment between the initial consonant and the main vowel, then we have no medials on the systematic phonetic level. They disappear in the process of realization, while the zero initial and the velar stops gain in their characterization.

3.4 Finals. Cantonese has a very rich final system. It has preserved the complete series of nasal finals: -m, -n, and -ŋ (the so-called yáng sheng yùn[ad] or yáng finals) and the complete series of stop finals: -p, -t, and -k (the so-called rù sheng yùn[ae] or rù finals). Words with vocalic finals (the so-called yin sheng yùn[af] or the yin finals) may end with almost any vowel or a vowel combined with -i or -u.

3.4.1 The yin finals. The class of yin finals consists of the main vowel optionally followed by another vocalic segment. This final vocalic segment may be -i, -u, or -ü in Cantonese. Since the occurrence of -ü can always be predicted from the preceding vowel, we shall represent both -u and -ü as /u/. We have a parallel in the Suzhou[ag] dialect of Wu, where [y] is always preceded by [ø].[5] The two final vocalic segments -i and -u correspond to Ancient Chinese *-i and *-u, as well as to the -i and -u in Mandarin, Hakka, Min, and Wu dialects, and may be looked upon as offglides. Here two supplementary rules are necessary:

$$\text{(a)}\quad [+\text{grave}] \longrightarrow [-\text{grave}] \;/\; \begin{bmatrix} -\text{cons} \\ +\text{voc} \\ -\text{grave} \\ +\text{flat} \\ -\text{diff} \end{bmatrix} \quad \begin{bmatrix} -\text{cons} \\ +\text{voc} \\ ____ \\ +\text{flat} \\ +\text{diff} \end{bmatrix} \quad \#$$

$$\text{(b)}\quad [\ \text{——}\ ] \longrightarrow [\text{-tense}] \;/\; \begin{bmatrix} \text{-cons} \\ \text{+voc} \\ \text{-grave} \\ \text{+flat} \\ \text{-diff} \end{bmatrix} \quad \begin{bmatrix} \text{-cons} \\ \text{+voc} \\ \text{-grave} \\ \text{-flat} \\ \text{+diff} \end{bmatrix} \quad \#$$

3.4.2 The rù finals, the rù tones, and syllable length. A rù final consists of a vowel followed by a stop consonant, -p, -t, or -k. The rù tones are ˥˥5, ˧˥3, ˨˩2, and ˧˥4. ˧˥4 is a colloquial variant of ˧˥3 and ˨˩2 and therefore not to be represented on the systematic phonemic level. It is generally believed that since rù tones correspond to rù finals, they are also shorter, because the final stop consonant abruptly checks the airstream. This is true only for certain vowels. The following is an inventory of rù finals which are as long as the yin and yáng finals:

	p	t	k
a	-ap	-at	-ak
e			-ɛk
ö			-ɯk
o		-ɔt	-ɔk
i	-ip	-it	
ü		-üt	
u		-ut	

Figure VII. Long rù finals.

Perfect correspondence of rù final, rù tone, and shorter duration is true only for the following rù finals:

	p	t	k
ə	-əp	-ət	-ək
ö		-öt	
i			-Ik
u			-Uk

Figure VIII. Short rù finals.

One possible explanation for the mismatch of stop finals with duration is that the main vowel overrides the stop ending in the determination of syllable length. We shall examine this notion in the next section.

3.4.3 Vowels and syllable length. It is claimed that there is a distinction between long and short vowels in Cantonese. Looking at the features in Figure V and the data in 3.4.2, we find a perfect correlation between the segmental feature [± tense] and the prosodic feature [± short]:

(a) [αtense] ⟶ [-αshort]

Therefore we have

i, ü, u, ɛ, ɑ, ɔ, a [+tense] and [-short]

I, U, e, ö, o, ə [-tense] and [+short]

Moreover, we find the following pairs: i ~ I, u ~ U, ɛ ~ e, and ɑ ~ ö. Members of each pair differ only in the characterization of the feature [± tense]. Again from the data in 3.4.2, we find that I is predictable from i and U is predictable from u by the following laxing rule:

(b) [+tense] ⟶ [-tense] / [-cons, +voc, αgrave, αflat, +diff] ___ [+cons, -son, +grave, -diff, -cont]

While ɛ and ɑ occur only before -ŋ and -k, they may be considered to result from their respective [- tense] counterparts by the following tensing rule:

(c) [-tense] ⟶ [+tense] / $\begin{bmatrix} -\text{cons} \\ +\text{voc} \\ -\text{grave} \\ -\text{diff} \\ ____ \end{bmatrix}$ $\begin{bmatrix} +\text{cons} \\ -\text{son} \\ +\text{grave} \\ -\text{diff} \end{bmatrix}$

Since the feature [± tense] is a major conditioning factor in determining syllable length, rules (b) and (c) should be ordered before rule (a).

The difference in vowel length for words with yin and yáng finals is not as obvious as with rù finals, since the voiced portion of the syllable is considerably lengthened and it is very difficult to separate the voiced portion of the main vowel from the voiced portion of its following [+vocalic] segment.

3.5 Reading (wén) vs. colloquial (bái) pronunciation. The following is a sample of words[ah] for which there is a distinction between reading and colloquial pronunciations:

Reading	Colloquial	Reading	Colloquial	Reading	Colloquial
tIŋ ˩ 22	tɛŋ ˩ 22	kIŋ ˥ 53	kɛŋ ˥ 53	tIk ˦ 3	tɛk ˦ 3
tsIŋ ˩ 22	tsɛŋ ˩ 22	lIŋ ˨˩ 21	lɛŋ ˨˩ 21	tsIk ˦ 3	tsɛk ˦ 3
mIŋ ˩ 22	mɛŋ ˩ 22	t'Iŋ ˥ 53	t'ɛŋ ˥ 53	sIk ˦ 3	sɛk ˦ 3
hIŋ ˥ 53	hɛŋ ˥ 53	t'sIŋ ˧˥ 35	t'sɛŋ ˧˥ 35	t'sIk ˦ 3	t'sɛk ˦ 3
t'sIŋ ˥ 53	t'sɛŋ ˥ 53	lIŋ ˦ 33	lɛŋ ˦ 33	p'Ik ˦ 3	p'ɛk ˦ 3

List A

Reading		Colloquial	
pun	22	p'un	23
tsɔ	22	t'sɔ	23
tün	22	t'ün	23
tam	22	t'am	23
kən	22	k'ə n	23

Reading		Colloquial	
kəu	33	k'əu	33
tiu	53	t'iu	53
t'sin	21	tsin	22
p'əŋ	21	pəŋ	22

List B

	Reading	Colloq
sai	33	35
kɛŋ	33	35
sɛ	33	35
jin	33	35

	Reading	Colloq
saŋ	53	55
pui	53	55
tin	35	55
kei	35	55

List C

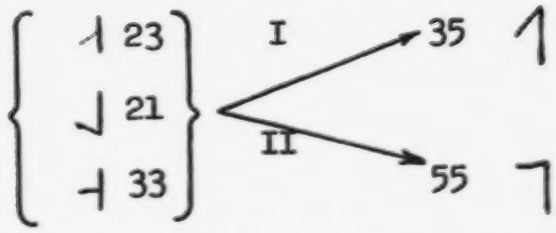
	Reading	Colloq
lɔ	21	35
t'iŋ	21	35
mun	21	35
t'sœŋ	21	35
p'un	21	35
mɛŋ	21	35
p'ɛŋ	21	35
sɛŋ	21	35
mən	21	55
mei	21	55
lUŋ	21	55

	Reading	Colloq
mou	23	35
lei	23	35
nai	23	55
laŋ	23	55
mei	23	55
təu	22	35
min	22	35
hɔŋ	22	35
lün	22	35
tsɔŋ	22	35
mui	22	35 55

List D

In the reading pronunciation of each of the words in List A, /i/ is the main vowel, which is realized as [I] by rule (c) of 3.4.3; in colloquial expressions the main vowel is /e/, which is realized as [ɛ] before ŋ and k. The tones are the same for both forms. For most of the words in List B, [-rising] → [+rising] $\left[\underline{\quad}\,{}^{-\text{high}}\right]$, while a change in aspiration takes place. Unaspirated expressions become aspirated in their colloquial forms. For the last two entries in the list, however, aspiration is dropped. This change is accompanied by a change in tone. For words in List C and List D, the difference in reading and colloquial pronunciations lies only in tone. These changes, though fairly regular, cannot be predicted by looking at the environments in terms of segmental or tonal features. In other words, they are to be marked on the systematic phonemic level by a feature which is neither segmental nor tonal. Therefore the feature [± colloquial] is introduced to bring about the necessary change.

A look at List D tells us that it is not a biunique mapping of tones between the reading and colloquial pronunciations. To solve this problem, let us assume that the following process takes place:



Another diacritic feature, [± colloquial'], is introduced to indicate the alternative process.

The following is a set of rules which accounts for all the reading and colloquial forms in the above lists except for the last two entries in List B.

(a) $[+\text{diff}] \longrightarrow [-\text{diff}] \;/\; \begin{bmatrix} -\text{cons} \\ +\text{voc} \\ -\text{grave} \\ -\text{flat} \\ -\text{tense} \\ \text{———} \end{bmatrix} \quad \begin{bmatrix} +\text{grave} \\ -\text{diff} \end{bmatrix}$

(b) $[-\text{tense}] \longrightarrow [+\text{tense}] \;/\; \# \begin{bmatrix} +\text{cons} \\ -\text{son} \\ -\text{cont} \\ +\text{colloq} \\ \text{———} \end{bmatrix}$

(c) $[-\text{contour}] \longrightarrow [+\text{rising}] \;/\; \begin{bmatrix} -\text{high} \\ +\text{colloq} \\ \text{———} \end{bmatrix}$

(d) $[-\text{high}] \longrightarrow \begin{bmatrix} +\text{high} \\ +\text{rising} \end{bmatrix} \;/\; \begin{bmatrix} +\text{colloq} \\ (+\text{colloq}') \\ \text{———} \end{bmatrix}$

(e) $[+\text{rising}] \longrightarrow [-\text{contour}] \;/\; \begin{bmatrix} +\text{high} \\ +\text{colloq}' \\ \text{———} \end{bmatrix}$

Though the last two entries in List B, t'sin ~ tsin and p'əŋ ~ pəŋ, may suggest a fairly reasonable phonological change, the change itself is not predictable. Moreover, the data are too limited to warrant any generalizations. Therefore, the reading form and the colloquial form would simply be entered as separate morphemes in our lexicon.

It seems that [− high] tones have a tendency to become [+ high] in their corresponding colloquial forms. Whereas

$$\left\{ \begin{array}{l} \text{˩ 21} \\ \text{˧ 23} \\ \text{˥ 22} \\ \text{˥ 33} \end{array} \right\} \longrightarrow \begin{array}{c} \text{˥} \\ 35 \end{array}$$

we have

$$\left\{ \begin{array}{l} \text{˥˧} \\ \text{˧˧} \end{array} \right\} \longrightarrow \text{˥˥}$$

for words with short rù finals. Examples are given in List E:

	Reading	Colloq
həp	˧˧	˥˥
kəp	˧˥	˥˥
löt	˧˧	˥˥
tsət	˧˧	˥˥
jUk	˧˧	˥˥
lUk	˧˧	˥˥

List E

The change of tone can be expressed by a simple rule:

$$\text{(f)} \quad [\text{-high}] \longrightarrow [\text{+high}] \; / \; \left[\begin{array}{l} \text{+colloq} \\ \text{+central} \\ \text{+short} \\ \underline{\qquad} \end{array} \right]$$

The colloquial forms of other rù finals which, by the nature of the main vowel, are [- short] can be generated by applying rules (a)-(d).

3.6 The alternation between the yin píng tones, ˥˥ and ˧˥. Of the three tonal alternations mentioned in 2.4, the first concerns a pair of yin píng tones. In an article which appeared in Zong Gúo Yǔ Wén[ak] 1964, Fú-bang Zong[al] held that, as a result of phonemic split, the two yin píng tones had separate phonemic status. However, a study of minimal pairs indicates that, in spite of semantic differences in contrastive environments, most of the cases are phonologically or morphologically conditioned. Bear in mind that, within the present theoretical framework, we are not so much concerned with the phonemic status of ˥˥ and ˧˥ as we are with the predictability of one tone from the other.

3.6.1 Phonological conditioning: tone sandhi. When preceding tone ˥55, ˧5, or ˨53, ˨53 becomes ˥55. So it is ˥55 that we always find in the following environment:

$$\longrightarrow \#\#\begin{Bmatrix} ˥55 \\ ˧5 \\ ˨53 \end{Bmatrix} \text{where \# stands for syllable boundary}$$

Our guess is that the comparatively lower ending pitch of ˨53 is assimilated to the relatively higher starting point of the pitch of the following syllable. For example:[am]

kou ˨53 + fUŋ ˥55 ⟶ kou fUŋ ˥55 ˥55

tsUŋ ˨53 + səm ˨53 ⟶ tsUŋ səm ˥55 ˨53

t'sɐŋ ˨53 + tsün ˨53 ⟶ t'sɐŋ tsün ˥55 ˨53

tsi ˨53 + tsUk ˧5 ⟶ tsi tsUk ˥55 ˧5

lɔk ˩22 + yIŋ ˨53 + pən ˨53 + fən ˨53 ⟶ lɔk yIŋ pən fən

˩22 ˥55 ˥55 ˨53

This tone sandhi can be expressed by the following rule:

$$[+\text{contour}] \longrightarrow [-\text{contour}] \;/\; \begin{bmatrix} +\text{high} \\ \text{____} \end{bmatrix} \#\, \# \begin{bmatrix} +\text{high} \\ -\text{rising} \\ -\text{central} \end{bmatrix}$$

3.6.2 Morphological conditioning. According to Fú-bang Zong, nearly seventy-five per cent of the Cantonese words with yin píng tones that he examined had only one tone value each. This was either ˥55 or ˨53, never both tones for the same character. This finding led Zong to believe that ˥55 and ˨53 could not be variants of the same tone. We are also told that, except for the tone sandhi mentioned above, each tone remained constant and stable in various phonological environments. Zong

nevertheless, made a very interesting observation. He found that a word with a yin píng tone almost always takes on the tone value ˥55 when it is used as a nominal, whether in isolation or in connection with another word. As nonnominals, a yin píng word is most likely of tone ˧53. This observation may allow us to mark the two yin píng tones alike on the systematic phonemic level and then to impose a morphological conditioning on the systematic phonemic representations to bring out the respective shapes.

Before we do that, let us examine the remaining twenty-five per cent of the data. Here each word has two possible tones, ˥55 and ˧53. However, these two yin píng tones are not in free variation with each other. Their occurrences are somehow predictable. Here are some further observations which may confirm our earlier guess of morphological conditioning.

(1) ˥55 cooccurs with nominals (N) when ˧53 cooccurs with verbs (V), adjectives (A), and measure words (MW).

(2) In nominal phrases (NP) consisting of a head (H) and a modifier (Mod) (i.e. NP → Mod + H), ˥55 goes with the head when ˧53 goes with the modifier, even though the modifier may be a nominal in its isolated form.

Let us generalize and express the alternation between ˧53 and ˥55 as a morphologically conditioned process:

$$\begin{bmatrix} +\text{high} \\ -\text{rising} \\ -\text{short} \end{bmatrix} \longrightarrow [-\text{contour}] \;/\; \left[X \,\#\, \begin{bmatrix} +\text{N} \\ +\text{count} \\ \text{———} \end{bmatrix} \,\#\, \text{N(H)} \right]_{\text{NP}}$$

X may be null

		in...	meaning...	used as a...
pou	˥55	fan pou ˧22 ˥55	rice-cooker	N
	˧53	pou fan ˧53 ˧22	to cook rice	V
sɔ	˥55	isolation ˥55	a comb	N
	˧53	isolation ˧53	to comb	V
hün	˥55	jün hün ˨21 ˥55	a circle	N
	˧53	hün wak ˧53 ˧22	to mark	V
t'an	˥55	t'sɔi t'an ˧33 ˥55	a vegetable stall	N
	˧53	fən t'an ˥55 ˧53	to snare, to split	V
tɛŋ	˥55	tɛŋ hai ˥55 ˩21	spikes	N
	˧53	tɛŋ hai ˧53 ˩21	to nail a shoe	V
tsəi	˥55	jœk tsəi ˧2 ˥55	medicine, drug	N
	˧53	jət tsəi jœk ˥5 ˧53 ˨2	a dose of medicine	MW
t'san	˥55	tsou t'san ˧35 ˥55	breakfast	N
	˧53	sam t'san (fan) ˥55 ˧53 (22)	three meals	MW
kəi	˥55	kUŋ kəi ˥55 ˥55	rooster	N (H)
	˧53	kəi tan ˧53 ˩35	egg	Mod
kUŋ	˥55	kei kUŋ ˧22 ˥55	technician	N (H)
	˧53	kUŋ t'sɔŋ ˧53 ˩35	factory	Mod
k^{w}ən	˥55	hɔi k^{w}ən ˩35 ˥55	navy	N (H)
	˧53	k^{w}ən hau ˧53 ˧22	military academy	Mod

$$\begin{bmatrix} +\text{high} \\ -\text{rising} \\ -\text{short} \end{bmatrix} \longrightarrow [+\text{contour}] \Big/ \left\{ \begin{bmatrix} \# & \begin{bmatrix} +\text{N} \\ +\text{count} \end{bmatrix}_{\text{Mod}} \# \# [+\text{N}] \# \\ & \qquad\qquad\qquad\qquad\quad \text{H} \end{bmatrix} \quad \begin{matrix} [_____]_{\text{V, A, MW}} \\ \\ \\ \\ \text{NP} \end{matrix} \right\}$$

This difference in the assignment of Cantonese yin píng tones reflects a difference in syntactic categories, a phenomenon which has a parallel in the placement of English primary stress. For example, in English we have 'convert, a nominal, and con'vert, a verb. The primary stress is placed on the first syllable in the nominal but on the second syllable for the verb form. Exceptions to the rules given above would simply have the feature [± contour] marked on the systematic phonemic level. Otherwise, the yin píng tone would be characterized on the systematic phonemic level by only two features, [−high] and [− rising], since [−short] and [± contour] can be generated by the rules.

Another interesting phenomenon in Cantonese is that words with yin píng tones always have the form ⌐55 when they are used as the last syllable of a feminine proper noun. When they are used in a masculine proper noun, however, there is no regularity. Here are some examples of proper nouns and their Cantonese pronunciations:[ao]

feminine proper noun	lei hæŋ k'ɐn	˩23	˥55	˥55
	hUŋ sün kiu	˩21	˥55	˥55
	fUŋ siu t'sIŋ	˩21	˧35	˥55
	lUk tsü	˩ 2	˥55	
	mui lan fɔŋ	˩21	˩21	˥55
masculine proper noun	tsæŋ fei	˥55	˥55	
	kIŋ ɔ	˥55	˦53	
	t'söü sæi t'sæŋ	˩21	˧33	˦53
	au yæŋ sæu	˦53	˩21	˥55
	k^{w}an hɔn hIŋ	˦53	˧33	˥55

List G

Here we have the rule:

$$[+\text{contour}] \longrightarrow [-\text{contour}] \;/\; \left[X \,\#\, \left[\begin{array}{c} +\text{high} \\ \hline \end{array} \right] \# \right]_{\text{(fem, proper)}}$$

X is not null

3.6.3 Semantic conditioning. For certain nominals, ˥55 and ˦53 are sometimes used to differentiate between two different meanings. In such cases, the ˥55 pronunciation often represents a later development of semantic content which reflects a cultural change initiated by or introduced into the speech community. For example:[ap]

		in...		meaning...
fɔŋ	55	jœk fɔŋ	˧ 3 ˥55	an instruction; a prescription
	53	tUŋ fɔŋ	˥55 ˨53	the eastern side or direction
tsœŋ	55	t'ou tsœŋ	˨21 ˥55	seal; insignia
	53	wən tsœŋ	˨21 ˨53	an essay; a literary composition
köü	55	t'sa köü	˨21 ˥55	a teahouse
	53	sən köü	˥55 ˨53	a new residence
ka	55	fa hɔk ka	˧33 ˧33 ˥55	a chemist
	53	ka t'Iŋ	˨53 ˨21	family
jin	55	jin tsai	˥55 ˥35	cigarette
	53	jin fɔ	˨53 ˥35	fireworks
tsU	55	t'ɔŋ t'ai tsUŋ	˨21 ˧22 ˥55	a posthumous title given to an emperor
	53	tsou tsUŋ	˥35 ˨53	ancestors

List H

Since there is no other formal characterization to differentiate the two sets of meanings, we would simply consider the two pronunciations as homophones and mark them as separate lexical entries. A certain loss of generality is to be expected in this treatment, but in doing it any other way there would be a danger of overgeneralization.

The following is a list of words with yin píng tones which, in combination with another word in larger constructions, can be pronounced either as ˥55 or as ˨53. There is absolutely nothing in the environment or in the semantic context to indicate why ˥55 is preferred in one combination, ˨53 in another. We shall mark them as they are in the lexicon since generalization is impossible.

	in...		meaning...
t'sɛ	tan t'sɛ	˥55 ˥55	bicycle
	hei t'sɛ	˧33 ˦53	automobile
	fɔ t'sɛ	˧35 ˦53	train
	ma t'sɛ	˨23 ˥55	carriage
$k^{w}a$	wɐŋ $k^{w}a$	˩21 ˥55	cucumber
	sai $k^{w}a$	˥55 ˦53	watermelon
	ha mɐt $k^{w}a$	˦53 ˧ 2 ˥55	honeydew
pau	sɨ pau	˥55 ˥55	satchel
	ma pau	˩21 ˦53	sack
tan	tan k^{w}ɵi	˥55 ˧35	monorail
	tan yan	˦53 ˨23	one-eyed
	tan jan p'ɐt ma	˦53 ˩21 ˥ 5 ˨23	single-handed

List I^{aq}

Chapter IV. Syllable structure

4.1 The canonical form of Cantonese syllables. Since most Chinese morphemes are monosyllabic, the discussion of morpheme structure is largely the discussion of syllable structure. A Cantonese syllable has the following canonical form:

$$C_1(V_1)V_2\left(\left\{\begin{matrix}V_3\\C_2\end{matrix}\right\}\right) \quad \text{or} \quad C(V)V\left(\left\{\begin{matrix}V\\C\end{matrix}\right\}\right)$$

C_1 stands for any initial consonant including the zero initial, which would be marked [+ cons]; C_2 stands for any of the six consonant endings: -p, -t, -k and -m, -n, -ŋ. V_2 is the main vowel. V_1 may be considered an onglide, V_3 an offglide. V_1 is what is traditionally called a medial; V_3, together with V_2, forms what is traditionally called the yin final. The canonical forms of Cantonese monosyllables are thus:

	I	II	III
1	C_1V_2	$C_1V_2V_3$	$C_1V_2C_2$
2	$C_1V_1V_2$	$C_1V_1V_2V_3$	$C_1V_1V_2C_2$

If we further break down C_2 into two classes: N, the nasal endings, and T, the stop consonants, the canonical forms of Cantonese monosyllables will be:

	I	II	III	IV
1	C_1V_2	$C_1V_2V_3$	C_1V_2N	C_1V_2T
2	$C_1V_1V_2$	$C_1V_1V_2V_3$	$C_1V_1V_2N$	$C_1V_1V_2T$

Examples are:[ar]

I		II		III		IV	
ma	'horse'	mai	'to buy'	man	'slow'	mat	'to wipe'
kua→[k^{w}a]	'melon'	kuai→[k^{w}ai]	'strange'	kuan→[k^{w}an]	'to close'	kuat→[k^{w}at]	'to scrape'

The reason for making C_1 an obligatory segment is that, on the systematic phonetic level, a syllable always begins with a consonant. That is to say, the zero initial is realized in its various forms. Instead of generating a segment from nothing, we shall assume the existence of a segment in the underlying form. Applying the phonological rules of 3.3, we eventually get rid of the medial in the systematic phonetic representation. Hence, the longest syllable on the systematic phonemic level consists of four segments, while the longest syllable on the systematic phonetic level consists of only three segments. The canonical forms we are considering here are what we would expect to find in the lexicon, i.e. the systematic phonemic representations.

4.2 Some general facts about Cantonese syllable structure. The canonical forms of Cantonese give us the following information about Cantonese syllable structure:

(1) Every syllable contains at least two segments, i.e. the initial and the main vowel.

(2) The longest vowel chain is a triphthong, $V_1V_2V_3$.

(3) There is no consonant cluster.

(4) The longest syllable contains four segments.

(5) Any consonant may occur as the initial, but only p, t, k, m, n, and ŋ may occur in the final position of a syllable.

(6) V_1 and V_3 can only be /i/ and /u/.

(7) The first segment after the syllable boundary must be consonantal.

(8) The second segment after the syllable boundary must be nonconsonantal. (A corollary of (3).)

(9) If the segment immediately preceding a syllable boundary is consonantal, then the segment before it must be nonconsonantal. (A corollary of (3).)

(10) A segment preceded and followed by a nonconsonantal segment must be nonconsonantal.

(11) If the segment immediately preceding a syllable boundary is consonantal, it must be one of the following (cf. (5)):

p [+cons, +son, -grave, +diff] t [+cons, -son, -grave] k [+cons, -son, +grave, -diff]

m [+cons, +son, +grave, +diff] n [+cons, +son, -grave] ŋ [+cons, +son, +grave, -diff]

4.3 The distributional pattern of Cantonese sounds. The canonical forms have shown us the permissible Cantonese monosyllable; now we have to find out what the actual syllables are and account for them. We find that Cantonese sounds have the distributional pattern recorded in Figures IX-XI. It is often said that Cantonese has 20 initials and 53 finals; the figures in themselves, however, are insignificant. We are interested in finding out how the features interact when placed in neighboring segments. Let us examine the finals to see how V_2 interacts with the last segment of the syllable.

Segment 4 / Segment 3	ɸ	-i	-u	(-U)	-m	-n	-ŋ	-p	-t	-k
a	a	ai	au		am	an	aŋ	ap	at	ak
ɔ		ɔi	ɔu		ɔm	ɔn	ɔŋ	ɔp	ɔt	ɔk
ɛ	ɛ						ɛŋ			ɛk
e		ei								
ɑ	ɑ						ɑŋ			ɑk
ɵ				ɵu		ɵn			ɵt	
ɔ	ɔ	ɔi				ɔn	ɔŋ		ɔt	ɔk
o			ou							
i	i		iu		im	in		ip	it	
I							Iŋ			Ik
u	u	ui				un			ut	
U							Uŋ			Uk
ü	ü					ün			üt	

Figure IX. Cantonese finals.

To add to the list of finals, there are two syllabic finals, m̩ and ŋ̩, which occur in isolation. We have not included them in our discussion of canonical forms. We may add the following syllable structure rule:

$$(1)\quad [+\text{seg}] \longrightarrow \begin{bmatrix} +\text{cons} \\ +\text{son} \\ +\text{voc} \end{bmatrix} \;/\; \#___\#$$

ṃ and ŋ̣ can be specified as:

ṃ [+diff] ŋ̣ [-diff]

Of the 2 x 51 = 102 possible combinations of medials with finals, only 28 actually occur. They are:

medial -i-	yin final	iʊ̈ŭ, iəi, iəu, iɛ, iai, (iɔ), (ia)
	yáng final	iəm, iən, iUŋ, iɤŋ, iœŋ, iʊ̈n
	rù final	iəp, iət, iUk, iœk
medial -u-	yin final	ua, uəi, uai, uɔ
	yáng final	uaŋ, uən, uɛn, uɔ̈ŋ, uIŋ, uəŋ
	rù final	uak, uət, uɛt, uɔk, uIk

Figure X. Medial-final combinations on the systematic phonemic level.

According to rules (1) and (2) of 3.2.1, -i- becomes j and -u- becomes w when preceded by the zero initial. When preceded by a velar stop, -u- first contributes to the secondary lip rounding of the velar stop and then drops, according to rules (5) and (6) of 3.3. Medial -i- never occurs with any initial other than the zero initial. Medial -u- occurs only with a velar stop initial or the zero initial. As a result, we have initials j, w, k^w and k^{lw} introduced into the system, while -i- and -u- systematically disappear. The distributional pattern would appear like this on the level of systematic phonetic representations:

j	yin final	jöü, jəi, jəu, jɤ, jai, (jɔ), (ja)
	yáng final	jəm, jən, jUŋ, jɤŋ, jɛɛŋ, jön, jam
	rù final	jəp, jət, jUk, jɛk, jɛɛk
w	yin final	wɤ, wəi, wai, wɔ
	yáng final	wəŋ, wən, wɑn, wɔŋ, wIŋ, waŋ
	rù final	wak, wət, wɑt, wɔk, wIk
k^w	yin final	k^wə, k^wəi, k^wai, k^wɔ
	yáng final	k^wəŋ, k^wən, k^wan, k^wɔŋ, k^wIŋ, k^waŋ
	rù final	k^wək, k^wət, k^wat, k^wɔk, k^wIk
k^{lw}	yin final	k^{lw}ə, k^{lw}əi, (k^{lw}ai)
	yáng final	k^{lw}əŋ, k^{lw}ən, k^{lw}ɔŋ
	rù final	(k^{lw}ək)

Figure XI. Medial-final combinations at the systematic phonetic level.

Adding the initials to the finals, we have a total distributional pattern as shown in Figure XII. Figure XI is incorporated in the pattern, with j, w, k^w, and k^{lw} considered initials. Hence the total distributional pattern is given in terms of systematic phonetic representations. Words in parentheses are colloquial expressions without written forms or without corresponding expressions in other major dialects of China.

4.4 Sequential constraints. A knowledge of the interactions of features within the syllable would be of great help to us in simplifying our lexical representation. This knowledge can be obtained by studying the distribution patterns given in Figures IX, X, and XI. Here are some of our observations expressible as syllable-structure rules.

	a	ai	au	am	an	aŋ	ap	at	ak	əi	əu	əm	ən	əŋ	əp	ət	ək	ɛ	ɛŋ	ɛk	ei
p	+	+	+		+			+	+	+			+	+		+	+		(+)		+
p'	+	+	+		+	+		+	+	+			+	+		+	+		(+)	(+)	+
m	+	+	+		+	+		+	+	+	+		+	+		+	+	+	(+)		+
f	+	+			+			+	(+)	+	+		+			+					+
t	+	+		+	+		+	+		+	+	+	+	+	+	+	+	+	(+)	(+)	+
t'	+	+		+	+		+	+		+	+	+	+	+	(+)				(+)	(+)	
n	+	+	+	+	+		+	+		(+)	+	+	+	+	+	(+)	(+)	+			+
l	+	+	+	+	+	+	+	+	+	+	+	+	(+)	+	+	+	+		(+)	(+)	+
ts	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	(+)	(+)	
t's	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+		+	(+)	(+)	
s	+	+	+	+	+	+	+	+		+	+	+	+	+	+	+	+	+	(+)	+	+
k	+	+	+	+	+	+	+		+	+	+	+	+	+	+	+		+	(+)		+
k'	+	+	+							+	+	+	+		+	+		+	(+)	+	+
ŋ	+	+	+	(+)	+	+	+			+	+	+	+	(+)	+	(+)	+				
h	+	+	+	+	+	+	+		+	+	+	+	+	+	+	+			(+)	+	+
ʔ	+	+	+	(+)	+	+	+	+	+	+	+	+	+	(+)	+	(+)					
j	+	(+)		(+)		(+)	(+)		(+)	+	+	+	+		+	+		+	(+)		
w	+	+			+	+		+	+	+			+	+		+					
k^w	+	+			+	+		+	+	+			+	+		+					
k'^w	+	(+)				+			(+)	+			+	+							

Figure XII.

The total distributional pattern

	œ	œŋ	œk	öü	ön	öt	ɔ	ɔi	ɔn	ɔŋ	ɔt	ɔk	ou	i	iu	im	in	ip	it	Iŋ	Ik
p							+			+		+	+		+		+		+	+	+
p'							+			+		+	+		+		+		+	+	+
m							+			+		+	+		+		+		+	+	+
f							+			+		+								(+)	
t		+	+	+	+		+	+		+		+	+	(+)	+	+	+	+	+	+	+
t'	+			+			+	+		+		+	+		+	+	+	+	+	+	+
n		+		+			+	+		+		+	+	(+)	+	+	+	+		+	+
l	(+)	+	+	+	+	+	+	+		+		+	+		+	+	+	+	+	+	+
ts		+	+	+	+	+	+	+		+		+	+	+	+	+	+	+	+	+	+
t's		+	+	+	+	+	+	+		+		+	+	+	+	+	+	+	+	+	+
s		+	+	+	+	+	+	+		+		+	+	+	+	+	+	+	+	+	+
k	+	+	+	+			+	+	+	+	+	+	+		+	+	+	+	+	+	+
k'		+	+	+				+		+		+			+	+	+		+	+	
p							+	+	+	+		+	+								
h	+	+		+			+	+	+	+	+	+	+		+	+	+	+	+	+	
ʔ							+	+	+			+	+								
j		+	+	+	+									+	+	+	+	+	+	+	+
w							+			+		+								+	+
k^w							+			+		+								+	+
k'^w										+											

Figure XII.

The total distributional pattern (cont.)

	u	ui	un	ut	Uŋ	Uk	ü	ün	üt	Ψ	ϧ
p		+	+	+	+	+					
p'		+	+	+	+	+					
m		+	+	+	+	+					
f	+	+	+	+	+	+					
t					+	+		+	+		
t'					+	+		+	+		
n					+			+			
l					+	+		+	+		
ts					+	+	+	+	+		
t's					+	+	+	+	+		
s					+	+	+	+	+		
k	+		+		+	+					
k'	+			+	+						
ŋ											
h					+	+		+	+		
ʔ					+	+					
j					+	+	+	+	+		
w	+	+	+	+							
k^w											
k'^w											

Figure XII.

The total distributional pattern (cont.)

A. Sequential constraints of VC sequences.

(1) A consonant segment immediately preceding the syllable boundary is always nonvocalic, discontinuous, and nonstrident (i.e. m, n, ŋ, p, t, or k):

$$[+\text{cons}] \longrightarrow \begin{bmatrix} -\text{voc} \\ -\text{cont} \\ -\text{strid} \end{bmatrix} \; / \; [\text{———}]\,\#$$

(2) If the segment preceding the syllable boundary is consonantal, grave, and nondiffuse (i.e. ŋ and k), the segment before it must be ɑ, ə, I, U, ɛ, œ, or ɔ, never ü, u, i, e, or o.

(3) If the segment preceding the syllable boundary is consonantal, grave, and diffuse (i.e. m and p), the segment before it must be ɑ, ə, or i.

(4) If the segment preceding the syllable boundary is consonantal and nongrave (i.e. n and t), the segment before it must be ä, ŋ, i, u, ü, ö, or ɔ. Except for ö and ə, all the possible vowels are tense.

(5) The two grave and nonflat vowels (i.e. ɑ and ə) can occur before any of the six consonant endings.

(6) Of the nongrave vowels, I, ɛ, and œ go only with -ŋ and -k; ü and ö go only with -n and -t. Of the grave vowels, U always goes with -ŋ and -k and u with -n and -t.

(7) ŋ never occurs with -m and -p; i never occurs with -ŋ and -k; e and o never occur before a consonant.

(8) The previous observations may suggest that we could further specify the consonantal endings, using the vowel as a conditioning factor. This is, however, not only implausible but, in fact, circular in

view of our previous analysis. Our earlier treatment of vowel length in 3.4.3 has motivated us to use consonant endings as the conditioning factor to predict the specification of the vowel as [± tense] and eventually that of the syllable as [± short]. To avoid circularity, we have to give a more complete specification of the consonantal endings; we cannot use the vowel features as conditioning factors. The loss of economy can be compensated for by a simpler vowel specification and the gain in generality with respect to syllable-length prediction. Hence, vis-à-vis the whole system, we may have even gained in economy.

(9) Using the consonantal endings as our conditioning factor, we have the following syllable-structure rules to account for observations (2)-(7):

(a) $\begin{bmatrix} \alpha\text{grave} \\ \alpha\text{flat} \\ \text{-diff} \end{bmatrix} \rightarrow \begin{bmatrix} \text{[-tense]} \\ \\ \\ \text{[+tense]} \end{bmatrix}_1 \; / \; [\text{——}] \; \begin{bmatrix} \begin{bmatrix} \text{+cons} \\ \text{+grave} \\ \text{-diff} \end{bmatrix} \\ \begin{bmatrix} \text{+cons} \\ \text{-grave} \end{bmatrix} \end{bmatrix}_1$

(b) $\begin{bmatrix} \text{-grave} \\ \text{-flat} \\ \text{+diff} \end{bmatrix} \rightarrow [\text{+tense}] \; / \; [\text{——}] \; \begin{bmatrix} \text{+cons} \\ \text{+grave} \\ \text{+diff} \end{bmatrix}$

(c) $\begin{bmatrix} \text{-grave} \\ \text{-diff} \end{bmatrix} \rightarrow [\text{+tense}] \; / \; [\text{——}] \; \begin{bmatrix} \text{+cons} \\ \text{+grave} \\ \text{-diff} \end{bmatrix}$

(d) $\begin{bmatrix} -\text{grave} \\ +\text{flat} \\ -\text{diff} \end{bmatrix} \longrightarrow [-\text{tense}] \;/\; [\text{———}] \begin{bmatrix} +\text{cons} \\ -\text{grave} \end{bmatrix}$

Feature specification of the VC sequences.

Here and in other feature specifications in this section, we shall further abbreviate the symbols representing the distinctive features as follows:

C	=	cons	=	consonantal
S	=	son	=	sonorant
V	=	voc	=	vocalic
G	=	grave		
D	=	diff	=	diffuse
F	=	flat		
Cp	=	comp	=	compact
Str	=	strid	=	strident
Ct	=	cont	=	continuant
T	=	tense		

am $\begin{bmatrix} +G \\ -F \\ +Cp \end{bmatrix} \begin{bmatrix} +C \\ +S \\ +G \\ +D \end{bmatrix}$ ap $\begin{bmatrix} +G \\ -F \\ +Cp \end{bmatrix} \begin{bmatrix} +C \\ -S \\ +G \\ +D \end{bmatrix}$

ɔm $\begin{bmatrix} +G \\ -F \\ -CP \end{bmatrix} \begin{bmatrix} +C \\ +S \\ +G \\ +D \end{bmatrix}$ ɔp $\begin{bmatrix} +G \\ -F \\ -Cp \end{bmatrix} \begin{bmatrix} +C \\ -S \\ +G \\ +D \end{bmatrix}$

	+C	-S	+G	+D	+C	-S	+G	+C	-S	+G	+C	-S	+G	+C	-S	+G	+C	-S	+G	+C	-S	+G	+C	-S	+G
	-G	-F	+D		+G	-F	+Cp	+G	-F	-Cp	-G	+F	-D	+G	+F	-D	-G	+F	+D	+G	+F	+D	-G	+F	+D
ip				at				at			öt			st			it			ut			üt		
im				an				an			ön			sn			in			un			ün		
	-G	-F	+D		+G	-F	+Cp	+G	-F	-Cp	-G	+F	-D	+G	+F	-D	-G	+F	+D	+G	+F	+D	-G	+F	+D
	+C	+S	+G	+D	+C	+S	-G	+C	+S	-G	+C	+S	-G	+C	+S	-G	+C	+S	-G	+C	+S	-G	+C	+S	-G

$$a_J \begin{bmatrix} +G \\ -F \\ +Cp \\ -D \end{bmatrix} \begin{bmatrix} +C \\ +S \\ +G \\ -D \end{bmatrix} \qquad a_k \begin{bmatrix} +G \\ -F \\ -Cp \\ -D \end{bmatrix} \begin{bmatrix} +C \\ -S \\ +G \\ -D \end{bmatrix}$$

$$e_J \begin{bmatrix} -G \\ -F \\ -D \\ +D \end{bmatrix} \begin{bmatrix} +C \\ +S \\ +G \\ -D \end{bmatrix} \qquad e_k \begin{bmatrix} -G \\ -F \\ +D \\ -D \end{bmatrix} \begin{bmatrix} +C \\ -S \\ +G \\ -D \end{bmatrix}$$

$$a_J \begin{bmatrix} +G \\ +F \\ -D \\ -D \end{bmatrix} \begin{bmatrix} +C \\ +S \\ +G \\ -D \end{bmatrix} \qquad a_k \begin{bmatrix} +G \\ +F \\ -D \\ -D \end{bmatrix} \begin{bmatrix} +C \\ -S \\ +G \\ -D \end{bmatrix}$$

$$c_J \begin{bmatrix} -G \\ +F \\ +S \\ -D \end{bmatrix} \begin{bmatrix} +C \\ +G \\ -D \\ -D \end{bmatrix} \qquad c_k \begin{bmatrix} -G \\ +F \\ -S \\ +D \end{bmatrix} \begin{bmatrix} +C \\ +G \\ -D \\ -D \end{bmatrix}$$

$$I_J \begin{bmatrix} -G \\ -F \\ +S \\ +D \end{bmatrix} \begin{bmatrix} +C \\ +G \\ -D \\ -D \end{bmatrix} \qquad I_k \begin{bmatrix} -G \\ -F \\ -S \\ +D \end{bmatrix} \begin{bmatrix} +C \\ +G \\ -D \\ -D \end{bmatrix}$$

Uŋ $\begin{bmatrix} +G \\ +F \\ +D \end{bmatrix} \begin{bmatrix} +C \\ +S \\ +G \\ -D \end{bmatrix}$ Uk $\begin{bmatrix} +G \\ +F \\ +D \end{bmatrix} \begin{bmatrix} +C \\ -S \\ +G \\ -D \end{bmatrix}$

B. Sequential constraints of diphthongs: VV/C___#.

(1) The segment following a consonant is always nonconsonantal and vocalic, as is a segment preceding a consonant.

$$[+seg] \longrightarrow \begin{bmatrix} -cons \\ +voc \end{bmatrix} \; / \; \left\{ \begin{matrix} [+cons]___ \\ ___ [+cons] \end{matrix} \right\}$$

(2) Any nonconsonantal segment in a noninitial position of the syllable is vocalic.

$$[-cons] \longrightarrow [+voc] \; / \; \#[+cons] \; X \; [\text{———}] \; Y \; \#$$

where X and/or Y may be null; the total number of segments contained in X and Y may not exceed 2.

(3) The second segment of a diphthong must be i, u, ɑ, ɔ, or ɛ. These are all tense vowels.

$$[\quad] \longrightarrow [+tense] \; / \; [+cons] \; [+seg] \; \begin{bmatrix} -cons \\ ___ \end{bmatrix} \; \#$$

(4) If the second segment of a diphthong is ɑ or ɔ, the first segment must be i or u. If the second segment is ɛ, the first segment must be i. In other words, if the second segment of a diphthong is nondiffuse, the first segment must be diffuse.

(a) $[-\text{cons}] \longrightarrow \begin{bmatrix} \alpha\text{grave} \\ \alpha\text{flat} \\ +\text{diff} \end{bmatrix} \; / \; [+\text{cons}] \; [\text{———}] \; \begin{bmatrix} -\text{cons} \\ +\text{grave} \end{bmatrix} \; \#$

(b) $[-\text{cons}] \longrightarrow \begin{bmatrix} -\text{grave} \\ -\text{flat} \\ +\text{diff} \end{bmatrix} \; / \; [+\text{cons}] \; [\text{———}] \; \begin{bmatrix} -\text{cons} \\ -\text{grave} \\ -\text{flat} \\ -\text{diff} \end{bmatrix} \; \#$

(c) $[-\text{cons}] \longrightarrow [+\text{diff}] \; / \; [+\text{cons}] \; [\text{———}] \; \begin{bmatrix} -\text{cons} \\ -\text{diff} \end{bmatrix} \; \#$

(5) If the second segment of a diphthong is u, and if the first segment is nondiffuse, it must be ö, o, ə, or ä. If the first segment is diffuse, it must be i.

(6) If the second segment of a diphthong is i, and if the first segment is nondiffuse, it must be e, ɔ, ə, or ä. If the first segment is diffuse, it must be u.

(7) From (5) and (6) we have

$[-\text{cons}] \longrightarrow \begin{bmatrix} -\alpha\text{grave} \\ -\alpha\text{flat} \end{bmatrix} \; / \; [+\text{cons}] \; [\text{———}] \; \begin{bmatrix} -\text{cons} \\ \alpha\text{grave} \\ \alpha\text{flat} \end{bmatrix} \; \#$

(8a) If i is preceded by a nongrave, flat, and nondiffuse vowel, it becomes rounded, i.e. ü, by assimilation:

$[-\text{flat}] \longrightarrow [+\text{flat}] \; / \; \begin{bmatrix} -\text{cons} \\ -\text{grave} \\ +\text{flat} \\ -\text{diff} \end{bmatrix} \; \begin{bmatrix} -\text{cons} \\ -\text{grave} \\ +\text{diff} \\ \text{———} \end{bmatrix} \; \#$

(8b) $\begin{bmatrix}-\text{grave}\\+\text{flat}\\+\text{diff}\end{bmatrix} \longrightarrow [-\text{tense}] \;/\; \begin{bmatrix}-\text{cons}\\ \underline{\quad}\end{bmatrix} \quad \begin{bmatrix}-\text{cons}\\-\text{grave}\\+\text{diff}\\+\text{flat}\end{bmatrix} \quad \#$

(9) If the first segment of a diphthong is $\begin{bmatrix}+\text{grave}\\+\text{flat}\\-\text{diff}\end{bmatrix}$ (i.e. o or ɔ), and if the second segment is diffuse, then the characterization of the feature [± tense] depends on the characterization of the feature [± grave] and [± flat] of the second segment:

$\begin{bmatrix}+\text{grave}\\+\text{flat}\\-\text{diff}\end{bmatrix} \longrightarrow [-\alpha\text{tense}] \;/\; [+\text{cons}] \quad \begin{bmatrix}-\text{cons}\\ \underline{\quad}\end{bmatrix} \quad \begin{bmatrix}-\text{cons}\\ \alpha\text{grave}\\ \alpha\text{flat}\\+\text{diff}\end{bmatrix} \quad \#$

(10) If the first segment of a diphthong is diffuse, it must be tense. If it is nondiffuse, it must be nontense. Therefore,

$[\alpha\text{diff}] \longrightarrow [\alpha\text{tense}] \;/\; [+\text{cons}] \quad \begin{bmatrix}-\text{cons}\\ \underline{\quad}\end{bmatrix} \quad [-\text{cons}] \quad \#$

(11) If the first segment of a diphthong is compact, it must be tense.

$[+\text{comp}] \longrightarrow [+\text{tense}] \;/\; [+\text{cons}] \quad \begin{bmatrix}-\text{cons}\\ \underline{\quad}\end{bmatrix} \quad [-\text{cons}] \quad \#$

(12) If the first segment of a diphthong is nondiffuse, the following segment must be diffuse:

$[\;\underline{\quad}\;] \longrightarrow [+\text{diff}] \;/\; [+\text{cons}] \quad \begin{bmatrix}-\text{cons}\\-\text{diff}\end{bmatrix} \quad \begin{matrix}-\text{cons}\\ \underline{\quad}\end{matrix}$

Feature specification of diphthongs: VV/C___#.

iu

+D

-C
+G
+F
+D

ui

+D

-C
-G
-F
+D

iɔ

-G
-F

-C
+G
+F

uɔ

+G
+F

-C
+G
+F

ia

-G
-F

-C
+G
-F

ua

+G
+F

-C
+G
-F

iɛ

+seg

-C
-G
-F
(-D)

ʊi

-G
+F
-D

-C
-G
+D

ɔi

+G
+F
-D

-C
-G
-F
(+D)

ou

+G
+F
-D

-C
+G
+F
(+D)

ai

+G
-F
+Cp

-C
-G
-F
(+D)

au

+G
-F
+Cp

-C
+G
+F
(+D)

ɔi $\begin{bmatrix} +G & -C \\ -F & -G \\ -Cp & -F \\ & (+D) \end{bmatrix}$ ɔu $\begin{bmatrix} +G & -C \\ -F & +G \\ -Cp & +F \\ & (+D) \end{bmatrix}$

ei $\begin{bmatrix} -G & -C \\ -D & -G \\ & -F \\ & (+D) \end{bmatrix}$

C. Sequential constraints of triphthongs: VVV.

(1) A triphthong is a combination of a medial and a yin final with two segments. The combinations of a medial and a yin final with one segment have already been treated in section B as diphthongs, together with yin finals consisting of two segments.

(2) Any nonconsonantal segment in a noninitial position of the syllable is vocalic.

(3) The first and the third of the nonconsonantal segments may be either i or u.

$$[\text{-cons}] \longrightarrow [\text{+diff}] \;/\; [\text{+cons}] \quad \left\{ \begin{matrix} [\text{——}] & [\text{-cons}] & [\text{-cons}] \\ [\text{-cons}] & [\text{-cons}] & [\text{——}] \end{matrix} \right\} \#$$

(4) If the first nonconsonantal segment of a triphthong is u, then the third must be i. If the third nonconsonantal segment of a triphthong is u, then the first must be i. In other words, we do not have a triphthong of the form *uVu in Cantonese.

(a) [-cons] ⟶ [-grave, -flat, +diff] / [+cons] [——] [-cons] [+grave, +flat] #

(b) [-cons] ⟶ [-grave, -flat, +diff] / [+cons] [+grave, +flat] [-cons] [____] #

(5) If the final segment of a triphthong is grave and flat (i.e. u), then the segment preceding it is [αgrave], [-αflat], nondiffuse and noncompact (i.e. ö or ə):

[-cons] ⟶ [αgrave, -αflat, -diff, -comp] / [+cons] [-cons] [——] [+grave, +flat] #

(6) If the final segment of a triphthong is nongrave and nonflat (i.e. i), then the segment preceding it is always grave, nonflat, and nondiffuse (i.e. ä or ə):

[-cons] ⟶ [+grave, -flat, -diff] / [+cons] [-cons] [——] [-grave, -flat] #

(7) Here again u becomes ü when preceded by ö.

Feature specification of triphthongs: VVV/C___#.

iəi [-G, -F] [+Cp] [-G, -F] iəi [-G, -F] [-Cp] [-G, -F]

uai	[+G, +F]	[+Cp]	[-C]		uəi	[+G, +F]	[-Cp]	[-C]
iöü	[-C]	[-G, +F]	[+G, +F]		iəu	[-C]	[+G, -F]	[+G, +F]

D. Sequential constraints of CVX sequences.

This can be subdivided into five cases:

1. where X = ∅; in this case we have a syllable #CV#.
2. where X = V; " " " " " " " #CVV#.
3. where X = VV; " " " " " " " #CVVV#.
4. where X = C; " " " " " " " #CVC#.
5. where X = VC; " " " " " " " #CVVC#.

In all these cases, the segment following the consonantal segment is nonconsonantal and vocalic. Now let us examine these cases one by one.

1. #CV#

(1) The segment following a consonantal segment and immediately preceding a syllable boundary is always nonconsonantal and tense.

[+seg] ⟶ [+tense] / [+cons] [——] #

(2) All tense vowels can occur in the position C_#. These tense vowels are ə, ɛ, œ, ɔ, i, u, and ü.

(3) Of the seven tense vowels that can occur in C_#, ə and ɔ have the widest distribution. ə occurs after a consonant, while ɔ occurs with all of the initial consonants but k'.

(4) We know from the discussion in 3.2.1 that if both ŋ and ʔ occur with the same subsequent sequence, only one of the two can be generated from the zero initial. In this case, the other would have to be fully specified.

(5) ɛ has a fairly wide distribution. It occurs after all except five of the initial consonants: t', p', f, l, h, and ŋ.

(6) i and ü occur only after strident consonants (i.e. ts, t's, and s) and the zero initial, which will eventually become the semivowel j according to rule (1) of 3.2.1. In colloquial expressions i also occurs after t and n, as in [ni ti] 'these'. The reading form of [ni] is [nei].

(7) u has a limited distribution. It occurs only after grave, [αdiffuse], [αcont] consonants (i.e. f, k, and k') and the zero initial, which according to rule (1) of 3.2.1 eventually becomes w.

(8) œ has the most limited distribution in C_#. It occurs in only three forms: [t'œ] (colloquial: [lœ]) 'to spit', [kœ] 'a saw', and [hœ] 'boat'.

(9) Not much can be said about vowels with a wide distribution. But for those with a more limited distribution in the environment C_#, we can make the following generalizations:

(a) $[+\text{cons}] \longrightarrow [+\text{strid}] \;/\; [\text{——}] \begin{bmatrix} -\text{grave} \\ +\text{diff} \end{bmatrix} \#$

(b) $[+\text{cons}] \longrightarrow \begin{bmatrix} +\text{grave} \\ \alpha\text{diff} \\ \alpha\text{cont} \end{bmatrix} \;/\; [\text{——}] \begin{bmatrix} +\text{grave} \\ +\text{diff} \end{bmatrix} \#$

(c) This is not so much a generalization as a statement of fact:

$$[+\text{cons}] \longrightarrow \left\{ \begin{bmatrix} +\text{grave} \\ -\text{diff} \end{bmatrix}, \begin{bmatrix} -\text{grave} \\ -\text{strid} \end{bmatrix} \right\} \;/\; [\text{——}] \begin{bmatrix} -\text{grave} \\ +\text{flat} \\ -\text{diff} \end{bmatrix} \#$$

Feature specification of #CV#. (Comprehensive List of Examples)

pa	-S +G +D -Ct -T	+G -F +Cp	pɔ	-S +G +D -Ct -T	+G +F -D	pɛ	-S +G +D (-Ct) -T	-G -F -D	t'œ	-S -G (-D) (+T)	-G +F -D			
p'a	-S +G +D -Ct +T	+G -F +Cp	p'ɔ	-S +G +D -Ct +T	+G +F -D	mɛ	+S +G +D	-G -F -D	lœ	+S -G (+V)	-G +F -D	[+colloq]		
ma	+S +G +D	+G -F +Cp	mɔ	+S +G +D	+G +F -D	tɛ	-S -G -D (-T)	-G -F -D	fu	+D +Ct	+G +F +D			
fa	-S +G +D +Ct	+G -F +Cp	fɔ	-S +G +D +Ct	+G +F -D	nɛ	+S -G (-V)	-G -F -D	ku	-D -Ct -T	+G +F +D			
ta	-S -G -D -T	+G -F +Cp	tɔ	-S -G -D -T	+G +F -D	tsɛ	-G +Str -Ct -T	-G -F -D	k'u	-D -Ct +T	+G +F +D			
t'a	-S -G -D +T	+G -F +Cp	t'ɔ	-S -G -D +T	+G +F -D	t'sɛ	-G +Str -Ct +T	-G -F -D	wu	+seg	+G +F +D			
na	+S -G -V	+G -F +Cp	nɔ	+S -G -V	+G +F -D	sɛ	-G +Str +Ct	-G -F -D	tsi	(-G) -Ct -T	-G -F +D			

la	+S -G +V		+G -F +Cp	lɔ	+S -G +V		+G +F -D	kɛ	+G -D -T (-Ct)		-G -F -D	t'si	(-G) -Ct +T		-G -F +D				
tsa	(-G) +Str -Ct -T		+G -F +Cp	tsɔ	(-G) +Str -Ct -T		+G +F -D	k'e	+G -D +T (-Ct)		-G -F -D	si	(-G) +Ct		-G -F +D				
t'sa	(-G) +Str -Ct +T		+G -F +Cp	t'sɔ	(-G) +Str -Ct +T		+G +F -D	kœ	+G -D -Ct (-T)		-G +F -D	ji	+seg		-G -F +D				
sa	(-G) +Str +Ct		+G -F +Cp	sɔ	(-G) +Str +Ct		+G +F -D	hœ	+G -D +Ct		-G +F -D	jü	+seg		-G +F +D				
tsü	(-G) -Ct -T		-G +F +D																
t'sü	(-G) -Ct +T		-G +F +D																
sü	(-G) +Ct		-G +F +D																

2. #CVV#

(1) In #CVV#, if the second segment is i or u, the first segment may be the zero initial, which would be realized as j or w by rule (1) of 3.2.1.

(2) ai and ɘi have the widest distribution. ai occurs with all initials and ɘi occurs with all but initial n.

(3) iu, ou, ɑu, and öu never occur after f.

(4) ɑu occurs after all other initials except t and t'.

(5) ou occurs after all other initials except k'.

(6) ɘu occurs after any initials but p and p'.

(7) iu, öu, and ei never occur after ŋ and ʔ.

(8) ei occurs after all other initials except t', ts, and t's.

(9) öu occurs after all other initials except p, p', m, and f.

(10) ɔi occurs after all initials except p, p', m, and f.

(11) Other than the zero initial, ui only occurs after k, k', p, p', m, and f, whereas uŋ occurs only after k, and ua only after k or k'.

(12) Here are some generalizations:

(a) $$[+\text{cons}] \rightarrow \left[\begin{bmatrix}+\text{grave}\\+\text{diff}\end{bmatrix}\begin{bmatrix}+\text{grave}\\-\text{diff}\\-\text{cont}\\-\text{tense}\end{bmatrix}\begin{bmatrix}+\text{grave}\\-\text{diff}\\-\text{cont}\end{bmatrix}\right]_1 \Big/ [\text{——}] \begin{bmatrix}+\text{grave}\\+\text{flat}\\+\text{diff}\end{bmatrix} \left[\begin{bmatrix}-\text{grave}\\-\text{flat}\\+\text{diff}\end{bmatrix}\begin{bmatrix}+\text{grave}\\+\text{flat}\\-\text{diff}\end{bmatrix}\begin{bmatrix}+\text{grave}\\-\text{flat}\\+\text{comp}\end{bmatrix}\right]_1 \#$$

(b) $$[+\text{seg}] \longrightarrow [+\text{cons}] \Big/ [\text{——}] \begin{bmatrix}\alpha\text{grave}\\\alpha\text{flat}\\+\text{diff}\end{bmatrix} [-\text{cons}] \#$$

(c) [+cons] $\left\{\begin{bmatrix}-\text{grave}\\-\text{diff}\end{bmatrix}\right.$ $\left.\begin{bmatrix}\text{grave}\\\text{diff}\end{bmatrix}\right\}$ / [——] $\begin{bmatrix}\alpha\text{grave}\\+\text{flat}\\-\text{diff}\end{bmatrix}$ $\begin{bmatrix}-\text{cons}\\-\alpha\text{grave}\\-\alpha\text{flat}\end{bmatrix}$ #

Feature specification of #CVV#. (Comprehensive List of Examples)

pai $\begin{bmatrix}-S\\+G\\+D\\-Ct\\-T\end{bmatrix}$ $\begin{bmatrix}+G\\-F\\+Cp\end{bmatrix}$ $\begin{bmatrix}-C\\-G\\-F\end{bmatrix}$ jiu $\begin{bmatrix}+seg\end{bmatrix}$ $\begin{bmatrix}+D\end{bmatrix}$ $\begin{bmatrix}-C\\+G\\+F\\+D\end{bmatrix}$ kui $\begin{bmatrix}-S\\+G\\-D\\-Ct\\-T\end{bmatrix}$ $\begin{bmatrix}+D\end{bmatrix}$ $\begin{bmatrix}-C\\-G\\-F\\+D\end{bmatrix}$

p'ai $\begin{bmatrix}-S\\+G\\+D\\-Ct\\+T\end{bmatrix}$ $\begin{bmatrix}+G\\-F\\+Cp\end{bmatrix}$ $\begin{bmatrix}-C\\-G\\-F\end{bmatrix}$ p'iu $\begin{bmatrix}-S\\+G\\+D\\-Ct\\+T\end{bmatrix}$ $\begin{bmatrix}+D\end{bmatrix}$ $\begin{bmatrix}-C\\+G\\+F\\+D\end{bmatrix}$ k'ui $\begin{bmatrix}-S\\+G\\-D\\-Ct\\+T\end{bmatrix}$ $\begin{bmatrix}+D\end{bmatrix}$ $\begin{bmatrix}-C\\-G\\-F\\+D\end{bmatrix}$

mai $\begin{bmatrix}+S\\+G\\+D\end{bmatrix}$ $\begin{bmatrix}+G\\-F\\+Cp\end{bmatrix}$ $\begin{bmatrix}-C\\-G\\-F\end{bmatrix}$ kou $\begin{bmatrix}-S\\+G\\-D\\-Ct\\-T\end{bmatrix}$ $\begin{bmatrix}+G\\+F\\-D\end{bmatrix}$ $\begin{bmatrix}-C\\+G\\+F\end{bmatrix}$

fəi $\begin{bmatrix}-S\\+G\\+D\\+Ct\end{bmatrix}$ $\begin{bmatrix}+G\\-F\\+Cp\end{bmatrix}$ $\begin{bmatrix}-C\\-G\\-F\end{bmatrix}$ hou $\begin{bmatrix}+G\\-D\\+Ct\end{bmatrix}$ $\begin{bmatrix}+G\\+F\\-D\end{bmatrix}$ $\begin{bmatrix}-C\\+G\\+F\end{bmatrix}$

təi $\begin{bmatrix}-S\\-G\\-D\\-T\end{bmatrix}$ $\begin{bmatrix}+G\\-F\\-Cp\end{bmatrix}$ $\begin{bmatrix}-C\\-G\\-F\end{bmatrix}$ məu $\begin{bmatrix}+S\\+G\\+D\end{bmatrix}$ $\begin{bmatrix}+G\\-F\\-Cp\end{bmatrix}$ $\begin{bmatrix}-C\\+G\\+F\end{bmatrix}$

t'əi $\begin{bmatrix}-S\\-G\\-D\\+T\end{bmatrix}$ $\begin{bmatrix}+G\\-F\\-Cp\end{bmatrix}$ $\begin{bmatrix}-C\\-G\\-F\end{bmatrix}$ fəu $\begin{bmatrix}-S\\+G\\-D\\+Ct\end{bmatrix}$ $\begin{bmatrix}+G\\-F\\-Cp\end{bmatrix}$ $\begin{bmatrix}-C\\+G\\+F\end{bmatrix}$

$$
\text{lai}\begin{bmatrix}+S\\ -G\\ +V\end{bmatrix}\begin{bmatrix}+G\\ -F\\ -Cp\end{bmatrix}\begin{bmatrix}-C\\ -G\\ -F\end{bmatrix}\text{tsau}\begin{bmatrix}+Str\\ -Ct\\ -T\end{bmatrix}\begin{bmatrix}+G\\ -F\\ +Cp\end{bmatrix}\begin{bmatrix}-C\\ +G\\ +F\end{bmatrix}
$$

$$
\text{kei}\begin{bmatrix}+G\\ -D\\ -T\\ -Ct\end{bmatrix}\begin{bmatrix}-G\\ -D\end{bmatrix}\begin{bmatrix}-C\\ -G\\ -F\end{bmatrix}\text{tsau}\begin{bmatrix}+Str\\ -Ct\\ +T\end{bmatrix}\begin{bmatrix}+G\\ -F\\ +Cp\end{bmatrix}\begin{bmatrix}-C\\ +G\\ +F\end{bmatrix}
$$

$$
\text{sei}\begin{bmatrix}+Str\\ \\ +Ct\end{bmatrix}\begin{bmatrix}-G\\ -D\end{bmatrix}\begin{bmatrix}-C\\ -G\\ -F\end{bmatrix}\text{wa}\begin{bmatrix}+seg\\ \\ \end{bmatrix}\begin{bmatrix}+G\\ +F\end{bmatrix}\begin{bmatrix}-C\\ +G\\ +F\end{bmatrix}
$$

$$
\text{tei}\begin{bmatrix}-S\\ -G\\ -D\\ -T\end{bmatrix}\begin{bmatrix}-G\\ +F\\ -D\end{bmatrix}\begin{bmatrix}-C\\ -G\\ -F\end{bmatrix}\text{wa}\begin{bmatrix}+seg\\ \\ \end{bmatrix}\begin{bmatrix}+G\\ +F\end{bmatrix}\begin{bmatrix}-C\\ +G\\ +F\end{bmatrix}
$$

$$
\text{ŋai}\begin{bmatrix}+S\\ +G\\ -D\end{bmatrix}\begin{bmatrix}+G\\ +F\\ -D\end{bmatrix}\begin{bmatrix}-C\\ -G\\ -F\end{bmatrix}\text{k'wa}\begin{bmatrix}-S\\ +G\\ -D\\ -Ct\\ -T\end{bmatrix}\begin{bmatrix}+G\\ +F\end{bmatrix}\begin{bmatrix}-C\\ +G\\ +F\end{bmatrix}
$$

$$
\text{hai}\begin{bmatrix}+G\\ -D\\ +Ct\end{bmatrix}\begin{bmatrix}+G\\ +F\\ -D\end{bmatrix}\begin{bmatrix}-C\\ -G\\ -F\end{bmatrix}\text{k'wa}\begin{bmatrix}-S\\ +G\\ -D\\ -Ct\\ -T\end{bmatrix}\begin{bmatrix}+G\\ +F\end{bmatrix}\begin{bmatrix}-C\\ +G\\ -F\end{bmatrix}
$$

$$
\text{tsôü}\begin{bmatrix}+Str\\ -Ct\\ -T\end{bmatrix}\begin{bmatrix}-G\\ +F\\ -D\end{bmatrix}\begin{bmatrix}-C\\ -G\\ +D\end{bmatrix}\text{k'wa}\begin{bmatrix}-S\\ +G\\ -D\\ -Ct\\ +T\end{bmatrix}\begin{bmatrix}+G\\ +F\end{bmatrix}\begin{bmatrix}-C\\ +G\\ -F\end{bmatrix}
$$

$$
\text{tôüi}\begin{bmatrix}-S\\ -G\\ -D\\ -T\end{bmatrix}\begin{bmatrix}-G\\ +F\\ -D\end{bmatrix}\begin{bmatrix}-C\\ -G\\ +D\end{bmatrix}\text{wui}\begin{bmatrix}+seg\\ \\ \end{bmatrix}\begin{bmatrix}+D\end{bmatrix}\begin{bmatrix}-C\\ -G\\ -F\\ +D\end{bmatrix}
$$

4. #CVC#

(1) Both -un and -ut occur with initials p, p', m, f, and the zero initial. Furthermore, -un occurs with k, and -ut occurs with k'.

$$[+\text{cons}] \longrightarrow \left(\begin{bmatrix} -\text{voc} \\ +\text{grave} \\ +\text{diff} \end{bmatrix} \left[\begin{bmatrix} -\text{son} \\ +\text{grave} \\ -\text{diff} \\ -\text{cont} \\ -\text{tense} \end{bmatrix} \begin{bmatrix} -\text{son} \\ -\text{grave} \\ -\text{diff} \\ -\text{cont} \\ +\text{tense} \end{bmatrix} \right]^{1} \right) \Big/ [\text{---}] \begin{bmatrix} +\text{grave} \\ +\text{flat} \\ +\text{diff} \end{bmatrix} \left[\begin{bmatrix} +\text{cons} \\ +\text{son} \\ -\text{grave} \end{bmatrix} \begin{bmatrix} +\text{cons} \\ -\text{son} \\ -\text{grave} \end{bmatrix} \right]^{1}$$

(2) Both -ön and -öt occur with initials ts, t's, s, and l. Furthermore, -ön occurs with initial t and the zero initial.

$$[+\text{cons}] \longrightarrow \left(\left[\begin{bmatrix} +\text{son} \\ -\text{grave} \\ +\text{voc} \end{bmatrix} \\ \emptyset \right]^{1} \begin{bmatrix} -\text{son} \\ -\text{grave} \\ +\text{strid} \end{bmatrix} \right) \Big/ [\text{---}] \begin{bmatrix} -\text{grave} \\ +\text{flat} \\ -\text{diff} \end{bmatrix} \left[\begin{bmatrix} +\text{cons} \\ +\text{son} \\ -\text{grave} \end{bmatrix} \begin{bmatrix} +\text{cons} \\ -\text{son} \\ -\text{grave} \end{bmatrix} \right]^{1}$$

(3) Both -ɔn and -ɔt occur with k and h. -ɔn also occurs with the zero initial which, according to rule (3) of 3.2.1, will be realized as ŋ. However, some pronunciations favor ʔ rather than ŋ. We shall have to specify ʔ in the lexicon.

$$[+\text{cons}] \longrightarrow \left\{ \begin{bmatrix} -\text{voc} \\ -\text{son} \end{bmatrix}, \begin{bmatrix} \emptyset \end{bmatrix}, \begin{bmatrix} -\text{son} \\ +\text{grave} \\ -\text{diff} \\ -\text{tense} \end{bmatrix} \right\}_1 \Big/ [\text{——}] \begin{bmatrix} +\text{grave} \\ +\text{flat} \\ -\text{diff} \end{bmatrix} \begin{bmatrix} \begin{bmatrix} -\text{cons} \\ +\text{son} \\ -\text{grave} \end{bmatrix} \\ \\ \begin{bmatrix} +\text{cons} \\ -\text{son} \\ -\text{grave} \end{bmatrix} \end{bmatrix}_1$$

(4) Both -ĩn and -ĩt occur with t, t', l, ts, t's, s, h, and the zero initial. -ĩn also occurs with n.

$$[+\text{cons}] \longrightarrow \left(\begin{bmatrix} -\text{son} \\ -\text{grave} \end{bmatrix}, \begin{bmatrix} +\text{son} \\ +\text{grave} \\ +\text{voc} \end{bmatrix}, \begin{bmatrix} -\text{son} \\ +\text{grave} \\ -\text{diff} \\ +\text{cont} \end{bmatrix}, \begin{bmatrix} \begin{bmatrix} +\text{son} \\ -\text{grave} \\ -\text{voc} \end{bmatrix} \\ \\ \begin{bmatrix} \emptyset \end{bmatrix} \end{bmatrix}_1 \right) \Big/ [\text{——}] \begin{bmatrix} -\text{grave} \\ +\text{flat} \\ +\text{diff} \end{bmatrix} \begin{bmatrix} \begin{bmatrix} +\text{cons} \\ +\text{son} \\ -\text{grave} \end{bmatrix} \\ \\ \begin{bmatrix} +\text{cons} \\ -\text{son} \\ -\text{grave} \end{bmatrix} \end{bmatrix}_1$$

(5) -ɐ̃n, -ɐ̃t, -ə̃n, -ə̃t, -in and -it have very wide distributions. -ɐ̃n, -ə̃n, and -ə̃t occur with all consonants (including the zero initial) except k'. Neither -in nor -it occur with ŋ. -it does not occur with n. -ə̃t occurs with all but four initial consonants: k, k', ŋ, and h.

Some generalizations:

(a) [+grave] → [+diff] / [+cons, -son, ___] [+grave, -flat, +comp] [+cons, -son, -grave]

(b) [+grave, -diff, -cont] → [-tense] / [+cons, -son, ___] [+grave, -flat] [+cons, -grave]

(c) [+son] → {[-grave], [+grave, +diff]} / [+cons, ___] [-grave, -flat, +diff] [+cons, -grave]

(d) [+son] → {[+grave], [-grave, +voc]} / [+cons, ___] [-grave, -flat, +diff] [+cons, -son, -grave]

(6) -ɐm, -ɐp, -əm, -əp, -im, and -ip never occur with initial p, p', m, or f. This distributional pattern shows signs of dissimilation.

[+cons] → {[-grave], [+grave, -diff]} / [___] [-flat] [+cons, +grave, +diff]

(7) -ɐm and -ɐp do not occur with k'. -ip does not occur with k' and ŋ. -im does not occur with -ŋ.

(a) $\begin{bmatrix}+\text{grave}\\-\text{diff}\\-\text{cont}\end{bmatrix} \rightarrow [-\text{tense}] \;/\; \begin{bmatrix}+\text{cons}\\-\text{son}\\ ___\end{bmatrix} \quad \left\{ \begin{bmatrix}+\text{grave}\\-\text{flat}\\+\text{comp}\end{bmatrix} \begin{bmatrix}+\text{cons}\\+\text{grave}\\+\text{diff}\end{bmatrix} \right.$
$\left. \begin{bmatrix}-\text{grave}\\-\text{flat}\\+\text{diff}\end{bmatrix} \begin{bmatrix}+\text{cons}\\-\text{son}\\+\text{grave}\\+\text{diff}\end{bmatrix} \right\}$

$\begin{bmatrix}+\text{grave}\\-\text{diff}\end{bmatrix} \rightarrow [-\text{son}] \;/\; \begin{bmatrix}+\text{cons}\\ ___\end{bmatrix} \quad \begin{bmatrix}-\text{grave}\\-\text{flat}\\+\text{diff}\end{bmatrix} \begin{bmatrix}+\text{cons}\\+\text{grave}\\+\text{diff}\end{bmatrix}$

(8) -ɔŋ and -ɔk occur with all consonants (including the zero *initial*). -ɔk also occurs with ʔ.

(9) -Uk does not occur with n and k'. -Uŋ and -Uk occur with all other consonants except ŋ.

(a) $[+\text{cons}] \rightarrow \left\{ \begin{bmatrix}-\text{son}\end{bmatrix}, \begin{bmatrix}+\text{son}\\-\text{grave}\end{bmatrix}, \begin{bmatrix}+\text{son}\\+\text{grave}\\+\text{diff}\end{bmatrix} \right\} \;/\; [___] \begin{bmatrix}+\text{grave}\\+\text{flat}\\+\text{diff}\end{bmatrix} \begin{bmatrix}+\text{cons}\\+\text{grave}\\-\text{diff}\end{bmatrix}$

(b) $\begin{bmatrix}-\text{son}\\+\text{grave}\\-\text{diff}\end{bmatrix} \rightarrow [-\text{tense}] \;/\; \begin{bmatrix}+\text{cons}\\ ___\end{bmatrix} \begin{bmatrix}+\text{grave}\\+\text{flat}\\+\text{diff}\end{bmatrix} \begin{bmatrix}+\text{cons}\\-\text{son}\\+\text{grave}\\-\text{diff}\end{bmatrix}$

(c) $[+\text{son}] \rightarrow \begin{bmatrix}+\text{grave}\\+\text{diff}\end{bmatrix} \;/\; \begin{bmatrix}+\text{cons}\\ ___\end{bmatrix} \begin{bmatrix}+\text{grave}\\+\text{flat}\\+\text{diff}\end{bmatrix} \begin{bmatrix}+\text{cons}\\-\text{son}\\+\text{grave}\\-\text{diff}\end{bmatrix}$

(10) Both -ɛŋ and -ɛk occur with t, l, ts, t's, s, k, k' and j. -ɛŋ also occurs with n and h.

(a) [+seg] → [-cons, -voc, +son, -grave] / [—] [-grave, +flat, -diff] [+cons, +grave, -diff]

(b) [+cons] → {[-grave, +strid], [-grave, -strid, -tense], [+grave, -diff]} / [—] [-grave, +flat, -diff] [+cons, +grave, -diff]

(c) [+cons, +son] → [-grave, -voc] / [—] [-grave, +flat, -diff] [+cons, +son, +grave, -diff]

(d) [+grave, -diff] → [-cont] / [—] [-grave, +flat, -diff] [+cons, -son, +grave, -diff]

(11) Both -ədŋ and -ədk occur with p', m, l, ts, t's, k, h, j, w, and the zero initial. -ədk also occurs with p, f, and ?. -ədŋ also occurs with s.

(a) [+cons, -son, +grave, +diff] → [+tense] / [—] [+grave, -flat, +comp] [+cons, +son, +grave, -diff]

(b) [+son] → {[+grave, +diff], [-grave, +voc]} / [+cons, ___] [+grave, -flat, +comp] [+cons, +son, +grave, -diff]

(c) $\begin{bmatrix} +\text{grave} \\ -\text{diff} \\ -\text{cont} \end{bmatrix} \rightarrow [-\text{tense}] \; / \; \begin{bmatrix} +\text{cons} \\ -\text{son} \\ ____ \end{bmatrix} \begin{bmatrix} +\text{grave} \\ -\text{flat} \\ +\text{comp} \end{bmatrix} \begin{bmatrix} +\text{cons} \\ +\text{grave} \\ -\text{diff} \end{bmatrix}$

In view of the fact that -əŋ and -ək do not occur with -k' either, (c) is modified to:

(c') $\begin{bmatrix} +\text{grave} \\ -\text{diff} \\ -\text{cont} \end{bmatrix} \rightarrow [-\text{tense}] \; / \; \begin{bmatrix} +\text{cons} \\ -\text{son} \\ ____ \end{bmatrix} \begin{bmatrix} +\text{grave} \\ -\text{flat} \end{bmatrix} \begin{bmatrix} +\text{cons} \\ +\text{grave} \\ -\text{diff} \end{bmatrix}$

(d) $[+\text{strid}] \rightarrow [-\text{cont}] \; / \; \begin{bmatrix} +\text{cons} \\ ____ \end{bmatrix} \begin{bmatrix} +\text{grave} \\ -\text{flat} \\ +\text{comp} \end{bmatrix} \begin{bmatrix} +\text{cons} \\ -\text{son} \\ +\text{grave} \\ -\text{diff} \end{bmatrix}$

(12) Both -əŋ and -ək occur with p, p', m, t, n, l, ts, s, and ŋ.

-əŋ also occurs with t', t's, k, h, ʔ, and w.

(a) $\begin{bmatrix} +\text{grave} \\ +\text{diff} \end{bmatrix} \rightarrow [-\text{cont}] \; / \; \begin{bmatrix} +\text{cons} \\ -\text{son} \\ ____ \end{bmatrix} \begin{bmatrix} +\text{grave} \\ -\text{flat} \\ -\text{comp} \end{bmatrix} \begin{bmatrix} +\text{cons} \\ +\text{grave} \\ -\text{diff} \end{bmatrix}$

(b) $\begin{bmatrix} -\text{grave} \\ -\text{strid} \end{bmatrix} \rightarrow [-\text{tense}] \; / \; \begin{bmatrix} +\text{cons} \\ -\text{son} \\ ____ \end{bmatrix} \begin{bmatrix} +\text{grave} \\ -\text{flat} \\ -\text{comp} \end{bmatrix} \begin{bmatrix} +\text{cons} \\ -\text{son} \\ +\text{grave} \\ -\text{diff} \end{bmatrix}$

(c) $[+\text{strid}] \rightarrow [-\text{tense}] \; / \; \begin{bmatrix} +\text{cons} \\ -\text{son} \\ ____ \end{bmatrix} \begin{bmatrix} +\text{grave} \\ -\text{flat} \\ -\text{comp} \end{bmatrix} \begin{bmatrix} +\text{cons} \\ -\text{son} \\ +\text{grave} \\ -\text{diff} \end{bmatrix}$

(c') $\begin{bmatrix} -\text{son} \\ -\text{grave} \end{bmatrix} \rightarrow [-\text{tense}] \; / \; \begin{bmatrix} +\text{cons} \\ ___ \end{bmatrix} \begin{bmatrix} +\text{grave} \\ -\text{flat} \\ -\text{comp} \end{bmatrix} \begin{bmatrix} +\text{cons} \\ -\text{son} \\ +\text{grave} \\ -\text{diff} \end{bmatrix}$

(13) Both -Iŋ and -Ik occur with p, p', m, t, t', n, l, ts, t's, s, k, w, and the zero initial. -Iŋ also occurs with f, k', and h.

(a) $\begin{bmatrix} +\text{grave} \\ +\text{diff} \end{bmatrix} \rightarrow [-\text{cont}] \; / \; \begin{bmatrix} +\text{cons} \\ -\text{son} \\ ___ \end{bmatrix} \begin{bmatrix} -\text{grave} \\ -\text{flat} \\ +\text{diff} \end{bmatrix} \begin{bmatrix} +\text{cons} \\ +\text{grave} \\ -\text{diff} \end{bmatrix}$

(b) $\begin{bmatrix} +\text{grave} \\ -\text{diff} \end{bmatrix} \rightarrow \begin{bmatrix} -\text{cont} \\ -\text{tense} \end{bmatrix} \; / \; \begin{bmatrix} +\text{cons} \\ -\text{son} \\ ___ \end{bmatrix} \begin{bmatrix} -\text{grave} \\ -\text{flat} \\ +\text{diff} \end{bmatrix} \begin{bmatrix} +\text{cons} \\ -\text{son} \\ +\text{grave} \\ -\text{diff} \end{bmatrix}$

(c) $\begin{bmatrix} -\text{cons} \\ -\text{voc} \end{bmatrix} \rightarrow [+\text{son}] \; / \; \begin{bmatrix} ___ \end{bmatrix} \begin{bmatrix} -\text{grave} \\ -\text{flat} \\ +\text{diff} \end{bmatrix} \begin{bmatrix} +\text{cons} \\ +\text{grave} \\ -\text{diff} \end{bmatrix}$

(14) Almost all occurences of -ɛŋ and -ɛk are in colloquial forms derived from corresponding -Iŋ and -Ik words. In such cases, the characterization of the colloquial form is identical with that of the reading form, except for an additional diacritic feature [± colloquial] or [± colloquial'].

(15) Not all reading forms have corresponding colloquial forms. Hence, the distribution of -ɛŋ and -ɛk is somewhat more limited than that of -Iŋ and -Ik.

(16) Though all words with -ɛŋ finals are colloquial expressions, not all words with -ɛk finals are colloquial forms. In

Cantonese, sIk 'to eat' and sɛk 'stone' contrast with each other. While -Ik does not occur with k' and h, -ɛk does.

(17) Here are some facts about the distribution of the -ɛk which occurs in noncolloquial forms:

(a) [+strid] → [+cont] / [+cons, -son, ___] [-grave, -flat, -diff] [+cons, -son, +grave, -diff]

(b) [+grave, -diff] → {[+cont], [-cont, +tense]} / [+cons, -son, ___] [-grave, -flat, -diff] [+cons, -son, +grave, -diff]

Feature specification of #CVC#. (Comprehensive List of Examples)

p'aŋ	+C, -S, +G, +D, -Ct	+G, -F, +Cp	+C, +S, +G, -D	sɛk	+C, -S, +Str, +Ct	-G, -F, -D	+C, -S, +G, -D	'stone'
maŋ	+C, +S, +G, +D	+G, -F, +Cp	+C, +S, +G, -D	sɛk	+colloquial [+C, -S, +Str, +Ct] [-G, -F, +D] [+C, -S, +G, -D]			'tin'
kap	+C, -S, +G, -D, -Ct	+G, -F, +Cp	+C, -S, +G, +D	hœŋ	+C, -S, +G, -D, +Ct	-G, +F, -D	+C, +S, +G, -D	
hap	+C, -S, +G, -D, +Ct	+G, -F, +Cp	+C, -S, +G, +D	kœk	+C, -S, +G, -D	-G, +F, -D	+C, -S, +G, -D	
jit	+C'	-G, -F, +D	+C, -S, -G	tön	+C, -S, -G, -Str	-G, +F, -D	+C, +S, -G	

ʒat	+C'	+G	+C		sötː	+C	-G	+C		jim	[+C']	-G	+C
		-F	-S			-S		+S				-F	+S
	-S	+Cp	-G			(-G)	+F	+G				+D	+G
						+Str							+D
t'sat	+C	+G	+C		kɔŋ	+C	+G	+C					
	-S					-S		+S					
	(-G)	-F	-S			+G	+F	+G		jün	[+C']	-G	+C
	+Str					-D						+F	+S
	-Ct	+Cp	-G			-Ct	+D	-D				+D	-G
	+T					-T							
k'an	+C	+G	+C		kɔn	+C	+G	+C		kUŋ	+C	+G	+C
	-S					-S		+S			-S		+S
	+G	-F	+S			+G	+F	+G			+G	+F	+G
	-D					-D					-D		-D
	-Ct	-Cp	-G			-Ct	+D	-D			-Ct	+D	
	+T												

paŋ	+C	+G	+C		mak	+C	+G	+C		tak	+C	+G	+C
	-S							-S			-S		-S
	+G	-F	+S			+S	-F	+G			-G	-F	+G
	+D										-Str		
	-T	-Cp	-G			+D	-Cp	-D				-Cp	-D

laŋ	+ colloquial										
	+C	-G	+S								
	+S	-F	+G								
	-G										
	+V	+D	-D								

	+C	-G	+S								
	-S										
	+G	-F	+G								
	-G										
	+V	+D	-D								

hɛŋ $\begin{bmatrix} \text{+ colloquial} \\ \begin{bmatrix} +C \\ -S \\ +G \\ -D \\ +Ct \end{bmatrix} \begin{bmatrix} -G \\ -F \\ +D \end{bmatrix} \begin{bmatrix} +S \\ +G \\ -D \end{bmatrix} \end{bmatrix}$ hIŋ $\begin{bmatrix} +C \\ -S \\ +G \\ -D \\ +Ct \end{bmatrix} \begin{bmatrix} -G \\ -F \\ +D \end{bmatrix} \begin{bmatrix} +S \\ +G \\ -D \end{bmatrix}$

5. #CVVC#

(1) In #CVVC# the segment immediately following the initial segment must be nonconsonantal, vocalic, α grave, α flat, and diffuse (i.e. i or u).

$$[+\text{seg}] \rightarrow \begin{bmatrix} -\text{cons} \\ +\text{voc} \\ \alpha\text{grave} \\ \alpha\text{flat} \\ +\text{diff} \end{bmatrix} / \quad [+\text{seg}] \; [\text{——}] \; [-\text{cons}] \; [+\text{seg}]$$

(2) If the segment immediately following a consonantal segment is grave and flat, the initial consonantal segment must be k, k', or the zero initial.

$$[+\text{cons}] \rightarrow \begin{bmatrix} -\text{son} \\ +\text{grave} \\ -\text{diff} \\ -\text{cont} \end{bmatrix} / \; [\text{——}] \begin{bmatrix} +\text{grave} \\ \\ +\text{flat} \end{bmatrix} \; [-\text{cons}] \; [+\text{seg}]$$

(3) If the segment immediately following a consonantal segment is nongrave and nonflat, the initial segment must be the zero initial.

$$[+\text{seg}] \rightarrow [+\text{cons}'] / \; [\text{——}] \begin{bmatrix} -\text{grave} \\ \\ -\text{flat} \end{bmatrix} \; [-\text{cons}] \; [+\text{seg}]$$

(4) -uIŋ and -uIk occur after k, but not k'. -uɔk, -uət, -uat, and -uan do not occur after k', either.

(a) $\begin{bmatrix}-\text{son}\\+\text{grave}\\-\text{diff}\\-\text{cont}\end{bmatrix} \rightarrow [-\text{tense}] \,/\, \begin{bmatrix}+\text{cons}\\ ____\end{bmatrix} \begin{bmatrix}+\text{grave}\\+\text{flat}\end{bmatrix} \begin{bmatrix}-\text{grave}\\-\text{flat}\\+\text{diff}\end{bmatrix} \begin{bmatrix}+\text{cons}\\+\text{grave}\\-\text{diff}\end{bmatrix}$

(b) $\begin{bmatrix}-\text{son}\\+\text{grave}\\-\text{diff}\\-\text{cont}\end{bmatrix} \rightarrow [-\text{tense}] \,/\, \begin{bmatrix}+\text{cons}\\ ____\end{bmatrix} \begin{bmatrix}+\text{grave}\\+\text{flat}\end{bmatrix} \left\{ \begin{bmatrix}+\text{grave}\\+\text{flat}\\-\text{diff}\end{bmatrix} \begin{bmatrix}+\text{cons}\\-\text{son}\\-\text{diff}\end{bmatrix} \begin{bmatrix}+\text{grave}\\-\text{flat}\end{bmatrix} \begin{bmatrix}+\text{cons}\\-\text{son}\\-\text{grave}\end{bmatrix} \begin{bmatrix}+\text{grave}\\-\text{flat}\\+\text{comp}\end{bmatrix} \begin{bmatrix}+\text{cons}\\+\text{son}\\-\text{grave}\end{bmatrix} \right\}$

(5) In the sequence VVC, -i- does not occur with -an, -at, -əŋ, -ək, -ɛk, -öt, -ɔn, -ɔŋ, -ɔt, -ɔk, -un, and -ut. Hence,

(a) $[+\text{grave}] \rightarrow [-\text{flat}] \,/\, [+\text{seg}] \begin{bmatrix}(-\text{cons})\\-\text{grave}\\-\text{flat}\end{bmatrix} \begin{bmatrix}(-\text{cons})\\ ____\end{bmatrix} \begin{bmatrix}+\text{cons}\\-\text{grave}\end{bmatrix}$

(b) $\begin{bmatrix}+\text{grave}\\-\text{flat}\end{bmatrix} \rightarrow [\alpha\text{comp}] \,/\, [+\text{seg}] \begin{bmatrix}(-\text{cons})\\-\text{grave}\\-\text{flat}\end{bmatrix} \begin{bmatrix}(-\text{cons})\\ ____\end{bmatrix} \begin{bmatrix}+\text{cons}\\\alpha\text{grave}\\-\alpha\text{diff}\end{bmatrix}$

(c) $\begin{bmatrix}+\text{grave}\\+\text{flat}\end{bmatrix} \rightarrow [+\text{diff}] \,/\, [+\text{seg}] \begin{bmatrix}(-\text{cons})\\-\text{grave}\\-\text{flat}\end{bmatrix} \begin{bmatrix}(-\text{cons})\\ ____\end{bmatrix} \begin{bmatrix}+\text{cons}\\+\text{grave}\\-\text{diff}\end{bmatrix}$

(d) $\begin{bmatrix}+\text{cons}\\-\text{grave}\end{bmatrix} \rightarrow [+\text{son}] \,/\, [+\text{seg}] \begin{bmatrix}(-\text{cons})\\-\text{grave}\\-\text{flat}\end{bmatrix} \begin{bmatrix}(-\text{cons})\\-\text{grave}\\+\text{flat}\\-\text{diff}\end{bmatrix} [____]$

(6) In the sequence VVC, -u- does not occur with -ɑ̈m, -ɑp, -əm, -əp, -im, -ip, -ön, -öt, -ɔn, -ɔt, -in, -it, -ək, -ɛŋ, -ɛk, -œŋ, -œk, -Uŋ, -Uk, ün, or -üt. Hence we have

(a) $[-\text{cons}] \rightarrow \begin{bmatrix} +\text{grave} \\ -\text{flat} \end{bmatrix} / [+\text{seg}] \begin{bmatrix} +\text{grave} \\ +\text{flat} \end{bmatrix} [\text{——}] \begin{bmatrix} +\text{cons} \\ -\text{grave} \end{bmatrix}$

(b) $[+\text{grave}] \rightarrow [-\text{diff}] / [+\text{seg}] \begin{bmatrix} +\text{grave} \\ +\text{flat} \end{bmatrix} [\text{——}] [+\text{cons}]$

(c) $[-\text{cons}] \rightarrow \left\{ \begin{bmatrix} +\text{grave} \\ -\text{flat} \end{bmatrix}, \begin{bmatrix} \alpha\text{grave} \\ \alpha\text{flat} \\ -\alpha\text{diff} \end{bmatrix} \right\} / [+\text{seg}] \begin{bmatrix} +\text{grave} \\ +\text{flat} \end{bmatrix} [\text{——}] \begin{bmatrix} +\text{cons} \\ +\text{grave} \end{bmatrix}$

(d) $\begin{bmatrix} +\text{cons} \\ +\text{grave} \\ -\text{diff} \end{bmatrix} \rightarrow [+\text{son}] / [+\text{seg}] \begin{bmatrix} +\text{grave} \\ +\text{flat} \end{bmatrix} \begin{bmatrix} +\text{grave} \\ -\text{flat} \\ -\text{comp} \end{bmatrix} [\text{——}]$

Feature specification of #CVVC#.

jən $[+C'] \begin{bmatrix} -G \\ -F \\ (+D) \end{bmatrix} \begin{bmatrix} +G \\ -F \\ -Cp \end{bmatrix} \begin{bmatrix} +C \\ +S \\ -G \end{bmatrix}$

wəŋ $[+C'] \begin{bmatrix} +G \\ +F \\ (+D) \end{bmatrix} \begin{bmatrix} +G \\ -F \\ -Cp \end{bmatrix} \begin{bmatrix} +C \\ +G \\ -D \end{bmatrix}$

jön $[+C'] \begin{bmatrix} -G \\ -F \\ (+D) \end{bmatrix} \begin{bmatrix} -G \\ +F \\ -D \end{bmatrix} \begin{bmatrix} +C \\ -S \\ -G \end{bmatrix}$

$$k^{w}_{jm} \quad \begin{bmatrix} +C \\ +G \\ -D \\ +T \end{bmatrix} \begin{bmatrix} +G \\ +F \\ (+D) \end{bmatrix} \begin{bmatrix} +G \\ +F \\ -D \end{bmatrix} \begin{bmatrix} +C \\ +G \\ +S \\ -D \end{bmatrix}$$

$$wjk \quad [+C'] \begin{bmatrix} +G \\ +F \\ (+D) \end{bmatrix} \begin{bmatrix} +G \\ +F \\ -D \end{bmatrix} \begin{bmatrix} +C \\ -S \\ +G \\ -D \end{bmatrix}$$

$$k^{w}_{at} \quad \begin{bmatrix} +C \\ +G \\ -D \end{bmatrix} \begin{bmatrix} +G \\ +F \\ (+D) \end{bmatrix} \begin{bmatrix} +G \\ -F \\ +Cp \end{bmatrix} \begin{bmatrix} +C \\ -S \\ -G \end{bmatrix}$$

$$wat \quad [+C'] \begin{bmatrix} +G \\ +F \\ (+D) \end{bmatrix} \begin{bmatrix} +G \\ -F \\ +Cp \end{bmatrix} \begin{bmatrix} +C \\ -S \\ -G \end{bmatrix}$$

$$jomj \quad [+C'] \begin{bmatrix} -G \\ -F \\ (+D) \end{bmatrix} \begin{bmatrix} -G \\ +F \\ -D \end{bmatrix} \begin{bmatrix} +C \\ -S \\ +G \\ -D \end{bmatrix}$$

References

Chao, Yuen Ren. 1947. Cantonese primer. Cambridge, Harvard University Press.

Cheng, Chi-yun. National atlas of China.

Chomsky, Noam. 1965. Aspects of the theory of syntax. Cambridge, M.I.T. Press.

Chomsky, Noam, and Morris Halle. The sound pattern of English. (To be published by Harper and Row.)

Halle, Morris. 1959. The sound pattern of Russian. The Hague, Mouton.

Jakobson, Roman, Gunnar Fant, and Morris Halle. 1952. Preliminaries to speech analysis. M.I.T. technical report 13.

Jakobson, Roman, and Morris Halle. 1956. Fundamentals of language. The Hague, Mouton.

Lǐ, Fang-kuei. 1937. Languages and dialects. Chinese yearbook. Shanghai, Commercial Press. Revised and expanded version to appear in Readings in Chinese linguistics, ed. by William S-Y. Wáng (University of Chicago Press).

Lin, Lien-hsen. 1964. Spoken language phenomena of the Chinese in Hong Kong. The Chung Chi journal 3.2:132-61.

Wáng, Lì. 1937. Hǎn yǔ shǐ gǎo. (Outline of Chinese historical phonetics.) Shanghai.

Wáng, William S-Y. 1967. Phonological features of tone. International journal of American linguistics 33.2:93-105.

Zong, Fú-bang. 1964. Guanyu Guǎngzhouhuà yinpíng diào de fenhuà wèntí. (On the split of the Yin-píng tone in Cantonese.) Zhong guó yǔ wén 132.376-89.

Footnotes

[1] This study was supported by National Science Foundation Grant GS1430.

[2] The five subgroups are: (1) Yùe hăi. Its area of distribution includes the basin of Xijiang and the delta of Zhujiang (Pearl River). (2) Sì yì. Spoken in the towns of Táishan, Xinhùi, Kaipíng, Enpíng, and their vicinity. (3) Gao-Léi. Spoken in Gaozhuo, Léizhou, and their vicinity. (4) Qin-Lián. Spoken in Qinzhou, Liánzhou (Hepu), and their vicinity. (5) Gùi nán. Spoken in the southern part of the Guangxi province. This area includes Wúzhou, Róngxiàn, Yùlín, and Bóbái.

[3] A set of distinctive features was also listed in Jakobson, Fant, and Halle (1952), Jakobson and Halle (1956), and Halle (1959). I believe, however, that the set coming out in Chomsky and Halle will be an improved version and the most sophisticated of all with respect to current phonological theory.

[4] For example, Lì Wáng (1937:125) and Lien-hsen Lin (1964:135). But this is probably the best these linguists can do within their theoretical framework of distinctive segments rather than distinctive features.

[5] [y] and [ɸ] are the IPA transcription. For Cantonese, we have chosen the symbol [ö] for IPA [ɸ] and the symbol [ü] for IPA [y]. The sounds symbolized by ɐ range over IPA ɐ, ʌ, and ɒ.

[a]Guǎngxi (Kwangsi) 廣西

[b]Guǎngdong (Kwangtung) 廣東

[c]Yùe 粵

[d]yáng 陽

[e]rù 入

[f]Yùe hǎi 粵海

[g]ní 泥

[h]niáng 娘

[i]lái 來

[j]Hakka 客家

[k]Mǐn 閩

[l]Wú 吳

[m]Nán-hǎi 南海

[n]Yǐng 影

[o]Yí 疑

[p]Yú 餘

[q]Yún 雲

[r]wéi 微

[s]Words beginning with n: 糯, 泥, 尼, 南, 男, 年, 諾, 難, 納, 孃, 寧, 女, 你, 紐, 扭, 娘, 稟, 農, 那, 挪, 膿, 匿, 膩, 溺.

Words beginning with l: 邏, 羅, 梨, 藍, 籃, 利, 李, 連, 落, 樂, 蘭, 立, 懶, 柳, 簍, 浪, 涼, 龍, 籠, 呂, 覽, 鋝, 歷, 礫.

[t]qiè-yùn 切韻

[u]•ŋiu 愚, a member of the Yí group

[v]•ʔiu 迂, a member of the Yǐng group

w. ŋan 岸, a member of the Yí group

x. ʔan 按, a member of the Yǐng group

y. ŋau 傲, a member of the Yí group

z. ʔau 奧, a member of the Yǐng group

ac kaikǒu hu 開口呼

ad yáng sheng yùn 陽聲韵

ae rù sheng yùn 入聲韵

af yin sheng yùn 陰聲韵

ag Suzhou 蘇州

ah Words of List A:

tIŋ ~ tɛŋ 定, tsIŋ ~ tsɛŋ 淨, mIŋ ~ mɛŋ 令, hIŋ ~ hɛŋ 輕, t'sIŋ ~ t'sɛŋ 青,
kIŋ ~ kɛŋ 驚, lIŋ ~ lɛŋ 靈, t'Iŋ ~ t'ɛŋ 聽, t'sIŋ ~ t'sɛŋ 請, lIŋ ~ lɛŋ 親,
tIk ~ tɛk 笛, tsIk ~ tsɛk 赤, sIk ~ sɛk 錫, t'sIk ~ t'sɛk 脊, p'Ik ~ p'ɛk 劈.

Words of List B:

pun ~ p'un 伴, tsɔ ~ t'sɔ 坐, tün ~ t'ün 斷, tam ~ t'am 淡, kən ~ k'ən 近,
kəu ~ k'əu 購, tiu ~ t'iu 雕, t'sin ~ tsin 纏, p'əŋ ~ pəŋ 凭.

Words of List C:

sai 婿, kɛŋ 鏡, sɛ 漏, jin 現, saŋ 生, pui 杯, tin 碘, kei 己.

Words of List D:

lɔ 螺, t'iŋ 亭, mun 門, t'sœŋ 腸, p'un 盆, mɛŋ 名, p'ɛŋ 瓶, sɛŋ 城,
mən 紋, mei 微, lUŋ 蓬, mou 母, lei 李, nai 奶, laŋ 冷, mei 尾,
təu 豆, min 面, hɔŋ 卷, lün 亂, tsɔŋ 撞, mui 妹.

ai Words of List E:

həp 盒, kəp 鴿, lɐt 栗, tsət 蛭, jUk 玉, lUk 鹿.

aj yin píng 陰平調

ak Zhong Gúo Yǔ Wén 中國語文. The title of the article is 關於廣州話陰平調的分化問題.

al Fú-bang Zhong 宗福邦

am The examples are: kou fUŋ高峰, tsUŋ sam忠心, t'sɐŋ tsün青磚, tsi tsUk知足, and lɔk yIŋ pan fan 落英繽紛.

an Words in List F are: pou煲, sɔ梳, hün圓, t'an攤, tɛŋ釘, tsɐi劑, t'san餐, kɐi雞, kUŋ工, k^{w}ɐn單; fan pou飯煲, pou fan煲飯, sɔ梳, jün hün圓圓, hün wak圓劇, t'sɔi t'an菜攤, fɐn t'an分攤, tɛŋ hai 釘鞋, jæk tsɐi藥劑, jɐt tsɐi jæk一劑藥, tsou t'san早餐, sam t'san (fan)三餐飯, kUŋ kɐi公雞, kɐi tan雞蛋, kɐi kUŋ技工, kUŋ t'sɔŋ工廠, hɔi k^{w}ɐn海單, k^{w}ɐn hau單校.

ao Words in List G are: lei hæŋ k^{w}ɐn李香君, hUŋ sün kiu洪宣嬌 fUŋ siu t'sIŋ馮小青, lUk tsü綠珠, mui lan fɔŋ梅蘭芳, tsæŋ fei張飛, kIŋ ɔ荊軻, t'sɔü sɐi t'sæŋ徐世昌, au yæŋ sau歐陽修, k^{w}an hɔn hIŋ關漢卿.

ap Words in List H are: jæk fɔŋ藥方, tUŋ fɔŋ東方, t'ou tsæŋ圖章, mɐn tsæŋ文章, t'sa kɔü茶居, sɐn kɔü新居, fa hɔk ka化學家, ka t'Iŋ家庭, jin tsɐi烟仔, jin fɔ烟火, t'ɔŋ t'ai tsUŋ唐太宗, tsou tsUŋ祖宗.

aq Words in List I are: tan t'sɛ單車, hei t'sɛ汽車, fɔ t'sɛ火車, ma t'sɛ馬車, wɔŋ k^{w}a黃瓜, sɐi k^{w}a西瓜, ha mɐt k^{w}a哈蜜瓜, sü pau書包, ma pau麻包, tan k^{w}ɐi單軌, tan ŋan單眼, tan jɐn pɐt ma 單人匹馬.

ar

The examples are: ma 馬, mai 買, man慢, mat抹, k"a 瓜, k"ai 怪, k"an 關, and k"at 刮. In these examples, the tone marks are omitted.

The Chinese characters for words in Footnote 2 are:

Yùe hǎi 粵海, Xijiang 西江, Zhujiang 珠江, Sì yì 四邑, Táishan 臺山, Xinhùi 新會, Kaipíng 開平, Enpíng 恩平, Gaozhou 高州, Léizhou 雷州, Qinzhou 欽州, Liánzhou 廉州, Hépǔ 合浦, Gao-Léi 高雷, Qin-Lian 欽廉, Gùi nán 桂南, Wúzhou 梧州, Róngxiàn 容縣, Yùlín 鬱林, Bóbái 博白.

S-Y. Wang 王士元

F. K. Li 李方桂

Li Wang 王力

Han Yu Shi Gao 漢語史稿

Lien-Hsen Lin 林蓮仙

香港中國人的語言現象概說

The Chung Chi Journal 崇基學報